AD

LABORATORY REPORT NO. 320

A COMPUTER BASED BIOMEDICAL INFORMATION SYSTEM

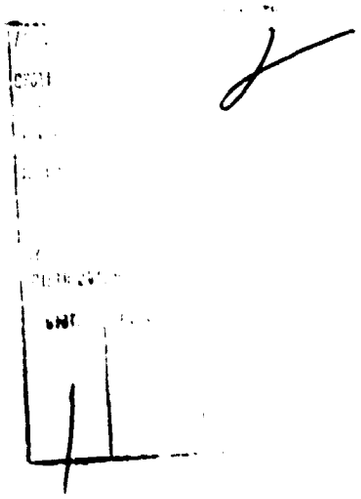
I. LOGIC FOUNDATION AND TECHNIQUES

by

JAMES C. SYNER
COMPUTER DIVISION

OCTOBER 1968

U. S. ARMY
MEDICAL RESEARCH &
NUTRITION LABORATORY
FITZSIMONS GENERAL HOSPITAL
DENVER COLORADO 80240



AD_____

U. S. ARMY MEDICAL RESEARCH AND NUTRITION LABORATORY

FITZSIMONS GENERAL HOSPITAL

DENVER, COLORADO 80240

LABORATORY
REPORT NO. 320

October 1968

Project Number: 3A025601A822 Military Internal Medicine

Task Number: 01 Biomedical Investigations

Work Unit No.: 067 Computer Classification of Pulmonary Disability

Project Number: 3A014501B71R Research in Biomedical Science

Task Number: 02 Internal Medicine

Work Unit No.: 064 Biomedical Information Systems Design

Project Number: 3A025601A822 Military Internal Medicine

Task Number: 00

Work Unit No.: 081 Digital Computer Based Biomedical Information System to Support Special Forces Troops in the Field

Project Number: 3A012501A03 Military Internal Medicine

Task Number: 01 Internal Medicine

Work Unit No.: 068 Computer Instrument Linkage

A COMPUTER BASED BIOMEDICAL INFORMATION SYSTEM

I. LOGIC FOUNDATIONS AND TECHNIQUES

James C. Syner, Colonel, M. C.

Computer Division

U. S. Army Medical Research and Nutrition Laboratory

PREFACE

The potential of the electronic digital computer in applications for medical practice and research has been presented with great enthusiasm and lofty expectations (1-14). In most writings the application of a digital computer to medical practice has been evaluated from the viewpoint of its functions as a general information processor rather than a high speed calculator. In the early writings the digital computer was advocated as an ideal answer to the dilemma in information processing confronting physicians and investigators.

The early articles of broad generality and high hopes have been followed by a sobering period of skepticism, disappointment and reassessment because of the failure to develop usable operations. However, those participating in the task of computerizing medical data and procedures have voiced a strong realization of the great complexity of a job made to appear straightforward and simple (15). Although computer performance in clinical medicine can attain the status of a fully integrated information system it must not be expected to be easily or quickly programmed.

At this time it is premature to speculate on points of broad generality regarding the impact of computers on medical practice. Until we have programmed biomedical information systems which permit the universal participation of physicians and investigators, the efforts to develop computer applications in medicine are basically research and development probes to confront and explore the problem.

TABLE OF CONTENTS

BODY OF REPORT

Work Unit Numbers:
- 067 Computer Classification of Pulmonary Disability
- 064 Biomedical Information Systems Design
- 081 A Digital Computer Based Biomedical Information System to Support Special Forces Troops in the Field
- 068 Computer Instrument Linkage

INTRODUCTION

An RCA 301 digital computer system was installed in December 1962 at the U. S. Army Medical Research and Nutrition Laboratory (USAMRNL) to meet the hardware requirements for a research project dedicated to study the problems involved in designing, developing and operating a computer based biomedical information system to service the needs of physicians engaged in patient care and clinical research, and scientists engaged in laboratory research. On the basis of long range planning it was projected that the results of this research study could serve as a basis for formulating recommendations to establish a large scale, real-time, on-line operational system for the Army Medical Service.

Establishing such a research project was viewed as a necessary approach to achieving systematic methods for handling the staggering problems in information processing confronting clinical physicians and laboratory investigators. The explosion of information which has occurred over the past three decades has resulted in data processing requirements totally beyond the response capabilities of manual systems. This is especially true for the data processing functions of data search, reduction, retrieval, analysis and display.

The installation of the digital computer system was the result of several years of effort (1957 - 1962) to improve the processing of biomedical information within the hospital-laboratory complex at Fitzsimons General Hospital. From the outset it was established as policy that the computer would be primarily programmed as a research tool and utilized in the same manner as any laboratory instrument. As segments of the programming achieved a fully debugged state they would be phased into selected operational runs at Fitzsimons General Hospital and the USAMRNL to test the concepts and methods. The term "operational runs" is used in a very restricted sense and refers to the controlled processing of data to satisfy test requirements. It does not refer to routine operational processing of records and reports on a full scale, round-the-clock basis at Fitzsimons General Hospital. The policy regarding service data processing is that only secondarily, based upon the availability of computer time and personnel,

would the division function as a service center for processing hospital and laboratory data on a sporadic basis outside of the programmed system.

A total biomedical information system for Fitzsimons General Hospital and the USAIRNL represents an enormously complex organization of components and functions. The development of an operational computer based system for such a complex will necessarily evolve through several generations of hardware-software alterations. There will probably never be a total solution to the problem, but rather sequential events viewed as minor to major achievements representing progress towards an ever improving system.

As progress is realized in attacking the problems of automation and systems development at Fitzsimons General Hospital it will become increasingly necessary to view the broader requirements of the Army Medical Service. The mission of the U. S. Army requires that troops perform duties on a world wide basis. Locating medical facilities and moving troops over widespread geographic areas are continuous phenomena within the system. To meet the requirements for medical services growing out of these circumstances attention must be concentrated on the technology of computer-communications to effect extensive networks linking medical facilities on a world wide basis. Ultimately a computer-communications network in which medical facilities at posts, camps and stations are linked to large scale data command centers at strategically positioned geographic locations will come to pass. This project is representative of the research efforts in systems development which must be implemented to realize the objective.

THE ROLE OF RESEARCH AND DEVELOPMENT

Accumulated experience indicates that evolving a complex biomedical computer facility presents a major challenge in research (16). In exploiting the application of computer technology to biomedical problems, provisions should be made for utilizing "Research and Development" notions, methods and facilities. It is a mistake to assume that biomedical applications for computers can be easily achieved merely through a skillful utilization of existing methods. One cannot ignore the possibility that the design of systems and the authoring of computer programs for biomedical applications may require an approach through research and development facilities.

The viewpoint that the development of a computerized biomedical information system offers a major challenge in research is based upon the following evidence:

1. Sufficient knowledge, systems development and operating experience are not currently available for this installation to support the implementation of a large biomedical computer system charged with providing immediate user services for such complex applications as patient care, delivery of health care services to large populations, medical research and medical education.

2. Borrowing logic and methods from other styled computer applications for direct utilization in biomedical matters is exceedingly vulnerable to failure if one is not prepared to investigate problems in systems analysis, systems design and computer programming.

3. The design of any computer based information system is a function of the environment within which it must operate, therefore, it is necessary to develop systems that are "tailor" made to the operational requirements of the environment of intended utilization.

THE SYSTEMS ANALYSIS

From the time of inception, and throughout its implementation, this project has been carried out under the concepts, principles and techniques of the systems analysis. The institution of this policy is motivated by the conviction that decisions regarding the structure and function of systems (automated, semi-automated, or manual) should be based upon the facts, and their evaluation, developed through the application of the principles and basic techniques of the systems analysis (17-20).

A definition and description of certain key terms and procedures are in order at this time to establish a basis for understanding the principles and techniques of systems analysis and design which have been adopted in the implementation of this research project. Although these terms and procedures are in common usage, and easily recognizable, they may have shades of meaning and slightly different connotations to different people. The key terms and proposed definitions are as follows:

1. Systems Analysis. The study in depth of the feasibility means and methods for accomplishing stated objectives with a view towards improving existing operations. Traditionally the systems study is documented in the form of a step-by-step logic flow diagram. Data Processing Systems Analysis is but one particular area in the total broad field of systems analysis and is concerned with the analysis of the fundamental phases in information processing, namely, collection, storage, retrieval, analysis and display.

2. Systems Design. The creation or invention of the new scheme for accomplishing stated objectives. The design is fashioned from the knowledge acquired during the systems analysis. Frequently the analysis and design of a system are regarded as a single task. However, it is advantageous to separate the two phases because the design of a new system leads to different actions from those encountered in the performance of the systems analysis. The systems design can be represented through the logic flow chart (figure 2). The flow chart technique shows that a system design involves four kinds of tasks:

A. Finding a set of functional blocks which can jointly perform all wanted functions.

B. Finding a set of devices to implement the functional blocks.

C. Making various kinds of computations and logical manipulations of the data.

D. Developing probability statistics on possible alternatives to assist in making the best choice. Expressing the system design in the flow chart provides a means for clearly identifying the information that must be available so that each step can be effectively executed.

3. The "System". In data processing systems analysis the achievement of a successful computer application for providing services to physicians and investigators is directly dependent upon the integration of all involved components and functions into a structured unit, the "System." As herein used a System is defined as:

A. The set of components and functions required to accept input, modify the input into output, and channel the output to its target of opportunity for performing some wanted operation on an object.

B. The components and functions assume a regular or orderly arrangement thereby forming a connected and interrelated series. Each component or function which is identified as a part of the whole process must be evaluated relative to all other components and functions to effect the best possible fit of its position and role to companion members. All necessary components and functions performed during each step in the life of a system can be represented in flow chart form.

The basic steps in the conduct of the systems analysis employed in this project included:

1. The study and analysis of the current system.

2. The identification and analysis of the requirements involved in designing a new system.

3. The identification and analysis of the requirements involved for operating and continuously updating the new system.

Once the systems analysis had provided a detailed documentation of the problem areas, activities were directed to:

1. Developing the logic which serves as a basis for selecting the components (hardware-software complex) of the new system.

2. Establishing the content and formats of the input and output documents.

3. Authoring the computer programs.

The practice of the systems analysis was not discontinued once the conceptual and early phases of the project were completed. The practice has been maintained throughout the operational phases of the project because of the vulnerability of the system to multiple forces, and the need to effect timely and appropriate adjustments. Any system (S-1) is continuously stressed by forces tending to throw it into a temporary or permanent state of malfunction (error commitment). Once the malfunction occurs a new system (S-2) exists, different from the original one by some finite degree. Under this condition of instability the user cannot realize his objectives, may stray considerably from his required course, and, unknowingly, initiate a new system (S-3). Therefore, a continuous practice of systems analysis is required to meet these circumstances and remain in command of the system and precisely oriented to the prescribed goals.

The system analysis has been conducted by the systems study team. The team membership has varied according to the problem under study. Separate teams have included mixtures of such specialists as physicians, mathematicians, statisticians, life scientists, engineers, pharmacists, laboratory technicians, nurses, administrators, computer programmers and systems analysts.

Through the systems analysis, facts were derived which made it possible to establish the following conditions of the system:

1. Feasibility.

2. Objectives.

3. Hardware requirements.

4. Software needs.

5. Personnel requirements.

6. Inputs.

7. Outputs.

PROJECT OBJECTIVES

The primary objective of the project is to design and program for the digital computer a prototype model of a biomedical information system for patient care and medical research at Fitzsimons General Hospital (FGH) and the USAMRNL. The achievement of this objective is relative and subject to the qualification that the need for refinement of the system in one way or another will be a continuing requirement. It is axiomatic that the "perfect" system is never developed, therefore, provisions must be made for continuing research and development activities to produce an improved system.

The secondary objective of the project is to test the designed system in a real situation of clinical practice, clinical research and laboratory technology. The scope of the operational processing will be limited to the requirements for fully testing the concepts and methods under operational demands. Initially the operational play will be implemented entirely under "batch" processing. In time, with refinement in logic programming and updating of the hardware complex, the operational play will evolve into a combination of "time-shared" and "batch" computer processing. The anticipation that the operational play will eventually be one of time-shared and batch processing is based upon existing evidence of the advantages and disadvantages of each in regard to such factors as:

1. Job cost.

2. Job elapse time.

3. Compiling time.

4. Operator requirements.

5. Programming time.

6. Turn-about time.

7. Availability of the central processing unit and I/O equipment for on-line debugging.

The final notions regarding optimum applications for time-shared and batch computer processing have not been established, and studies for evaluating these methods are in progress (22-27).

The future objective of the project is to interact with other computer groups within the Army Medical Service to participate in the design and development of an integrated computer-communications network that links computerized information systems based in Army medical facilities through the United States and overseas. The structure of such a computer-communications network is visualized as an integrated logic system of satellite data terminals and small computers linked to large scale data command centers located in strategic geographic regions.

BIOMEDICAL APPLICATIONS

In its fully developed form the research project will be involved with studies which include the following classes of biomedical applications for the digital computer: 1). Patient Care, 2). Health Care Services to Military Personnel on Active Duty, 3). Medical Research and 4). Medical Education. The pursuit of studies in these areas is formalized in protocols designed to attack the problems and obtain results and experience which can serve as a basis for the design and development of definitive operational systems.

1. Patient Care. The study of systems requirements to automate the professional and technical information involved in providing patient care services will cover inpatients and outpatients. The study of this application is primarily implemented through the protocol, "Computer Classification of Pulmonary Disability" (Work Unit Number 067). This protocol is regarded as the basic approach to the entire project. Principal supplementation for the study of this application is realized through the protocol, "Biomedical Information Systems Design" (Work Unit Number 064).

2. Health Care Services to Military Personnel on Active Duty.

The study of systems requirements to automate the information processing required to support the delivery of health care services to military personnel not in formal medical channels (i.e., military personnel

actually performing active duty) will involve the total health record. The study of this application is implemented through the pulmonary disability protocol (Work Unit No. 067), with major supplementation from the protocol, "A Digital Computer Based Biomedical Information System to Support Special Forces Troops in the Field" (Work Unit Number 081). This application must be operational within the "field" or "work-live" environment irrespective of geographic location. This notion is unique from the patient care sector (item 1) which is operational only within the relatively restricted environment of the fixed medical installation (hospital, clinic, dispensary). In exploring this application, the organizational and operational characteristics of the United States Army Special Forces, along with the training and deployment of its medical personnel, provides a unique target of opportunity for establishing a prototype model of a computer based biomedical information system to support troops in the field. The characteristics of personnel training, worldwide deployment, in-country operations, exposure to naturally occurring diseases, and a detailed documentation of pre- and post-mission examinations establish a sound basis for a study uniquely oriented to military requirements under field conditions. Some of the major problem areas which will be considered within the domain of this application include:

1. Immunization prophylaxis.

2. Epidemiologic mapping.

3. Nutrition and energy expenditure.

4. Environmental hostilities.

5. Physical and psychological training.

6. Pharmacologic factors in performance.

7. Troop education.

3. Medical Research. The study of computer applications in support of medical research is accomplished to some degree in all of the protocols authored under this research project. Because of the high frequency utilization in medical research of instrumentation for data acquisition, the protocol, "Computer Instrument Linkage" (Work Unit Number 068) is of particular value to the study of this application. This project has been established to study the programming and systems planning necessary to effect the automatic storage, retrieval, analysis and display of the digital output of an analog to digital conversion system attached to various clinical and laboratory

instruments. Through this study programming logic required to handle the digital output of a variety of analog instruments utilized routinely in clinical studies performed at FGH and research studies performed at USAMRNL will be achieved. This effort is a vital supplement to the pulmonary disease protocol (Work Unit Number 067). In analog to digital conversion systems wherein the volume or complexity of analysis of the output data is dependent on a digital computer, the absence of necessary computer programming is considered to be the single greatest obstacle to developing automatic systems.

4. Medical Education. A project to study the utilization of computer assisted instruction (CAI) to support programs in continuing medical education for the staff, residents and interns has not been designed at this time. It is anticipated that this application will be explored through the use of a special purpose hardware configuration consisting of on-line video data terminal linkage between medical personnel and the computer. The user's interrogations (asking the question) will be utilized as the principal catalyst to the unfolding of logic branching designed to permit a continuous movement in the direction of the correct answer. All kinds of graphics will be utilized in accomplishing data reduction to display information to the user (28, 29). The graphics can be generated automatically under programmed controls in response to the user's questions and commands. The video data terminal will consist of a CRT (cathode ray tube) controlled with the availability of electronic "pens" by user interaction with a given display on the CRT. An additional feature of the terminal will be a conventional keyboard augmented by a set of special function buttons in the user's console. The usual print facilities for quickly generating hard copies of displayed pictures will be included. The development of a software structure for such a system will present many complex problems. From a function point of view, the software must provide a linkage between a high level symbolic picture description and the CRT machine language display file through a set of generation subroutines. Current hardware-software techniques for displaying computer-generated alphanumeric and graphical information on a CRT display are powerful and offer a valuable tool to service user demands for continuing medical education.

APPROACH TO THE PROBLEM

The computer based biomedical information system which this project aims to develop will be an organization of enormous complexity in its completed form. The systems analysis established that the totality of records, functions, personnel actions and hardware involved in the system were far too enormous and complex to permit

the development of a computer based system through a single effort or crash program. Additionally, it was established that this program would be conducted under conditions of significant limitations in professional staff, knowledge, experience and physical resources. Accordingly, it was concluded that the approach of choice would be to design and develop a prototype or model system to simulate the characteristics of the "Grand System," and yet permit testing at such a reduced level that controls could be maintained easily and continuously.

A model of a total biomedical information system was established through the design of an exercise in the area of pulmonary diseases. This small scale model, "The Classification of Disability in Patients with Pulmonary Diseases," (Work Unit Number 067) provides the investigators the opportunity to simulate all aspects of the "grand system" and yet enjoy the simplicity, convenience and controls afforded by a special purpose model. The small system encompasses all aspects of information processing -- i.e., collections, storage, retrieval, analysis and presentation -- which would be available to a user in the "grand system."

As a model of an information system for delivering health care services, the classification disability resulting from pulmonary disease is viewed as a member of a class of medical problems. It was reasoned that a digital computer system for any one member of the class (i.e., gastrointestinal, cardiovascular-renal, musculoskeletal, neuropsychiatric, genitourinary, etc.) could serve as a model for the techniques, procedures, and hardware required to implement information processing for total patient care.

In attempting to simulate all of the steps involved in classifying disability in patients with pulmonary diseases and bank the data elements required to test the classification technique, problems arose which could not be handled within the design of the "disability" protocol. Either these problems had not been anticipated or their complexity had been underestimated. Throughout the implementation of the project these problems continued to arise and complicate the efforts to computerize the process of disability classification. Ultimately it became apparent that it would be necessary to take definitive steps to deal with these complications.

As the problems were analyzed and requirements for their solution itemized, considerations were given to the various means of resolving the complications. It was concluded that the procedure of choice was to supplement the "disability" protocol by compartmentalizing the problems into broad categories and designing additional protocols to cover each area. This approach was adopted in preference to aggregating the totality of problems into the single protocol.

Adopting this approach permits one to view the project as an interacting group of smaller projects (subsystems) that can be linked together around a single objective. After repeated reassessment of the problems which arose in attempting to automate the model, the following problem categories were established as supplements to the "disability" protocol:

1. Computer-instrument linkage.

2. Computer-communications through remote control send/receive data terminals.

3. Biomedical subsystems (clinics, wards, laboratories, pharmacy, services, etc.) in a hospital complex.

The supplements to the "disability" protocol were never intended to destroy its role as the principal approach to automating the processing of biomedical information. The supplements (subsystems) serve to identify major activities which must be integrated into the total system in order to achieve simulation of total patient care for any class of medical problem. The study and programming of these activities as though they were independent entities was considered to be a much simpler task than would be the case if they were aggregated into a single, enormously complex whole.

HARDWARE COMPONENTS

The hardware components can be divided into two broad categories referred to as "on-line" and "off-line." The components that are wired to the computer are referred to as being "on-line." These components operate automatically and sequentially under the direct control of programmed instructions executed in the computer high speed memory. The components that are not wired to the computer and do not operate automatically under computer programmed controls are referred to as "off-line."

The "on-line" components include:

1. RCA Model 305 Processor (computer). The processor is a general purpose, digital, stored program, transistorized machine. Within this unit is housed the following integrated units:

   A. Console.

   B. High Speed Memory.

   C. Program Control.

D. Input-Output Control Modules.

E. Power Supply.

The console panel provides for complete monitoring of the operation of the computer. The high speed memory is a random access, magnetic core device which provides storage and work area for programs and data. The memory has a capacity of 40,000 character locations, and each location is individually addressable and can store one of the RCA 301 characters. The inventory of RCA 301 characters includes all the letters of the alphabet, the ten decimal digits, control and special symbols. The memory cycle time to address, bring into a register, and regenerate a character in its original memory location is 4.7 microseconds. The program control executes the instructions of the program stored in the high speed memory and performs the automatic accuracy checks. The classes of instructions manipulated by the program control includes:

A. Data Handling.

B. Arithmetic Operations.

C. Decision and Control.

D. Input-Output.

2. RCA Model 321 Paper Tape Reader-Punch. This unit is mounted on one base and can read and punch at a rate of 100 characters per second. The packing density is 10 characters per inch, and the tape speed is approximately 10 inches per second. The reader stops on a character and is positioned to read the next character.

3. RCA Model 308 Monitor Printer. This typewriter-like device is operated under program control and prints up to 10 characters per second. All RCA 301 characters may be printed. The monitor printer is extremely useful in providing on-line communications between the core resident program and the operating personnel. It permits the operating personnel to produce a printed copy of the program execution. The monitor printer can be operated up to a distance of fifteen feet from the processor.

4. RCA Model 361 Data Record File. This is a random access bulk storage device. Up to six data record files may be used in the system. The data record file contains 128, 6 and 7/8ths inch diameter magnetic-coated discs. Storage of information is achieved by encoding data on both sides of the discs. The information is

either written to the disc or read from the disc to high speed memory, magnetically. Each surface of every record has two bands, each containing 10 cells of 900 characters each. Information is recorded in serial fashion in a spiral pattern around the record, and characters can be transferred from the data record file to computer memory in blocks of from one to ten cells at a rate of 2500 characters per second. Up to 4 and 1/2 million characters can be stored in one complete record file, and records can be manually removed and interchanged to increase the capacity. This is a random access device in that each or any band of a record may be directly addressed. This is directly opposed to the limited system of magnetic tape where the entire tape must be scanned to locate a given character or item.

5. RCA Model 382-4 Hi-Data Tape Group. This component provides a means for serial access (or storage) of data on magnetic tapes. The model 382 utilizes a one-half inch wide, Mylar-based magnetic tape. The tape is transported at 60 inches per second in a forward or reverse direction and 120 inches per second during rewind. The model 382 provides two modes of writing: 1). The 382 mode which operates at 30 KC, records at a density of 500 characters per inch, and gives the nominal read and write rate of approximately 30,000 characters per second; 2). The 381 mode which operates at 20 KC, records at a density of 333 characters per inch, and gives the nominal read and write rate of approximately 20,000 characters per second. The tape reel has a maximum diameter of eight inches and accommodates 1230 feet of magnetic tape of which a minimum of 1200 feet is available for recording data.

6. RCA Model 333 On-Line Printer. This is a high-speed data output device designed to print 120 columns of output data on single or multiple sheet fanfold paper. All electronics, print, and paper fed mechanisms are housed in one console cabinet. The printer is equipped with paper feed malfunction devises which detect a low paper or torn-paper condition. Paper advance is controlled by the print and paper advance instruction of the program. Paper advance may be on a line-by-line basis or by means of a format loop on the printer.

7. RCA 330-78 (IBM 1402) Card Read/Punch. This on-line I/O device reads 800 cards per minute and punches 250 cards per minute. It handles both Hollerith and binary codes in the read/punch modes. The device is equipped with an advanced read feature which achieves an additional 10 milliseconds of computer processing time during a card read operation. The device has the Stack/Select feature which permits one to select and separate certain cards, based upon imposed criteria, during a card punch operation. The cards selected are shunted from the main stream into card hoppers where they are sequentially stacked.

The off-line components include:

1. Friden Flexowriters and Printers. This equipment provides a paper tape punch and reader for preparing and editing tapes for computer input. Punched paper tape produced as computer output can be read off-line with the resultant production of a printed copy of the information. A Selecta-data unit coupled to the flexowriter provides a means for preparing hard copy and tape formats through mechanized programmatic controls. A verifier unit coupled to the flexowriter provides a means for performing an automatic edit of the punch paper tape.

2. Tally Model 311 Data Terminal, Transmit/Receive. This equipment provides transmission and receipt of digital data over dial up telephone at speeds of 1200 words per minute. The system will operate completely unattended, detects and deletes errors, and can be used for off-line tape duplication and editing. The equipment provides fully automatic controls over its functions of information processing through its capabilities to answer incoming calls, verify that an authorized transmitter or receiver is calling, and thence send or receive the data in punched paper tape format. At completion, it hangs up the call and turns itself off. These automatic features permit the computer division to handle communications during nonworking hours and have the data processed by the following morning.

3. Tape and Record Storage Cabinets. These storage units are required to safeguard program tapes and temperature-humidity sensitive tapes and discs when not in use. Backup files for programs and data are maintained in an air conditioned room in a building remote from the computer site where an image file of programs and tapes is also maintained.

INFORMATION COMPONENTS AND FUNCTIONS

As herein used the term "software" includes, 1). information files, and 2). library of computer program instructions. In conventional usage the term "software" is limited to mean "computer programmed instructions." The computer programmed instructions are designed and authored by man and express his will and intelligence in directing how the computer must handle the data. The logical similarities in the roles played by "data" and "instructions" in computer execution provides a basis for linking them together in an inclusive category of "software" components. So linked together, these components establish the information content of the system which is to be handled by the hardware components.

The software components of the biomedical information system can be categorized as follows:

1. Medical Knowledge. This includes all the traditional and generally accepted body of information of the life sciences specialty subjects such as physiology, pathology, chemistry, bacteriology, virology, pharmacology, anatomy, etc.

2. Patient Data. This includes the items of data collected on each patient. The data are representative of all sections of the traditional clinical case record such as chief complaint, history of present illness, past history, physical examination, laboratory data, data from special procedures, review of systems and therapy.

3. Computer Programmed Instructions. This includes all the logic schemes for data handling and analysis whereby input is modified into the desired output. The programs can be written in any one of the three following languages: 1). Machine or natural language, 2). Assembly language, and 3). Compiler language such as Fortran or the University of Miami Algebraic Language (UMAL).

The functional or action capabilities of the system consist of:

1. Responding to the user's interogations.

2. Accepting multiple data inputs.

3. Modifying input to produce output.

4. Effecting the transmission of information from laboratory and clinical instruments to the digital computer through analog-digital conversion systems.

5. Providing for remote control data terminal send/receive services through data terminal-computer linkage.

The functions listed above are programmed through the full range of information processing which includes:

1. Information Collection.

2. Information Storage.

3. Information Retrieval.

4. Information Analysis.

5. Information Display and Presentation.

LOGIC FOUNDATIONS OF THE INFORMATION SYSTEM

A schematic of the information system is presented in Figure I. As diagrammed, it is an accurate, though overly simplified, generalization of the system. The flow of information processing illustrated in Figure I demonstrates that "input," modified by "program commands," produces "output" which is intended to effect some wanted action on a physical object. Within the domain of this biomedical information system, the supreme action is the delivery of services that will effect an improvement in the health state of a particular human being or group of human beings.

Each of the major sectors -- "Input," "Computer," "Output" -- can in turn be amplified by branching diagrams depicting a hierarchical structure. This scheme illustrates the "molecular" structure of the system, and graphically portrays that the development of any increasingly complex system is achieved by an ordered and progressive aggregation of basic units or modules. It is axiomatic that any variety of a system, simple or enormously complex, is nothing more than an ordered aggregation of basic entities or modules.

Each sector represented in the schematic has a major functional role in the system. These can be categorized as follows:

1. Sector I, The Input: "To State the Problem." The nature and volume of the input are determined by the problem to be engaged. Once a statement of the problem has been formalized it is represented by compiled medical knowledge, patient data, and user's interrogations. As regards this latter component, the process of "asking the question" is considered to be one of the most important, and difficult to achieve, components of the input. The user's question is regarded as the catalyst for effecting a meaningful execution of computer information processing.

2. Sector II, The Computer: "To Respond to Problem Solving Requirements." When all pertinent "input" and programming have been written to the computer memory, this high speed electronic machine responds by executing the prescribed commands. The execution of the countless steps in logic and computation represents a delegated (given through man's will and intelligence) response to effect a solution to the problem. Many of the medical problems which are now forced on us by the complexities of a growing civilization could not be solved without computers. The total response to these problem solving requirements is a beautiful demonstration of collaboration between man and machine.

3. Sector III, The Output: "To Deliver and Display the Product." The completion of the information processing involves display of

the "output" on any one of several different devices such as printers, video terminals, or "X-Y" plotters. The display of the products of information processing to the user is vital to the ultimate delivery of the wanted operation. This function is being greatly improved through the newer developments in graphic displays which dramatically illustrate the idiom, "a good picture is worth a thousand words."

BIOMEDICAL COMPUTER CENTER

A schematic of the projected biomedical computer center is illustrated in Figure II. This flow-diagram represents the structure, functions, and information-flow pathways for the information processing complex. The complex is based upon the concepts of computer-communications over a network structure of stations and relays to meet the everyday requirements of varying data loads and processing complexities.

The complex is divided into four sectors with their attending components and integrating pathways. As illustrated in the schematic the sectors, components, and pathways appear highly discrete, independent and rigorously integrated along sure, firm pathways. In actuality there is considerable overlapping in the boundaries of the sectors and in the zones of interaction between the components. The information-flow pathways are subject to considerable variance as regards both direction and intensity of data movement. However, the principal pathways of data flow, the alignment of functions to components and the boundaries of the sectors are accurately represented in the schematic diagram.

The four sectors represented are designed to operate as follows:

1. The Biomedical Environment. This is the setting (ward, clinic, laboratory, office, etc.) where interpersonnel relationships are established between the medical staff and patients to effect:

   A. Problem Definition.

   B. Data Acquisition.

   C. Interrogation.

   D. Implementation of the Diagnosis-Treatment Complex.

2. The Clinical Assistance Computer. This is the site of the small-to-moderate scale computer complex where the following services are accomplished:

A. Maintenance and daily updating of the patient data files which are compiled from multiple input sources, such as: the ward, nurse's station, doctor's office, laboratory, pharmacy, specialty clinics, etc.

1). As patients are discharged from hospital or outpatient follow-up the clinical records are deleted from the active files of this sector of the system and stored in the files of the large scale computer processing sector of the system. When a patient is readmitted to hospital or the outpatient service, the past record can be accessed and placed on an active status in the files of the small scale (clinical assistance) computer.

B. Preparing print-outs of the daily medical notes, progress reports, clinical messages, and data tabulations for the medical staff. The various print-outs can be obtained under automatic programmed controls on interrogation by the user. This information service can be programmed for display on a video (CRT) terminal which can be equipped with electronic (light) pen and typewriter for on-line interrogation, data input, and data deletion-correction.

3. The Higher Level Processing Computer. This is the site of the large scale computer complex where the following functions are accomplished:

A. Maintenance of updated banks of medical and scientific information. The specialized data bases of medical and scientific knowledge are maintained in a current status through programs of information transfer and sharing with other institutions and military installations.

B. Maintenance of both current and historic files of patient data. These files are maintained for indefinite periods of time and provide large bases of data for population, statistical, and probabilistic studies on medical diagnosis-treatment problems.

C. Providing the physician and research scientist a complete battery of logic schemes for the manipulation and computation of data. These procedures pertain directly to decision making functions and consist of the broad logical functions of: statistical analysis, propositional calculus (symbolic logic), systematic classification (numerical taxonomy), probability theory, value theory, queueing theory, mathematical models, simulation etc.

D. Back-up assistance to the clinical assistance computer during periods of over-load.

E. Total information processing and computational support to the clinical research program. A very direct relationship exists between the capability to perform clinical research on clinical case records and the availability of a hospital communication system. The clinical investigator who attempts any sizable study on clinical case records through manual means faces a task of such enormous demands -- tedious, time consuming, unclassified information scattered over many records and pages, illegible scrip, etc. -- that very few physicians carry out this exceedingly important, vitally necessary work.

F. The performance of research studies on systems design, programming techniques, problem oriented languages, mathematical models, and statistical methods are delegated to this computer complex.

4. The Remote Control Data Terminal. This sector will meet the requirements for data send/receive operations between post, camps and stations and the large scale digital computer complex. This capability is absolutely essential to meeting the design criteria of the system which establish computer-communication operations. An approach to the study, development and evaluation of remote control data transmission/ receiving is formalized under two protocols, namely, Work Unit No. 067, (Study I), and Work Unit No. 081. The implementation of these protocols is in progress. A schematic of the computer-communication operations for the initial phase of Work Unit No. 067, Study I, is presented in Figure III. This schematic is also applicable for the initial phase of actions in Work Unit No. 081. The immediate goals of these projects are as follows:

A. To establish the capability for transmitting and receiving medical data from remote stations under automatic controls.

B. To link the terminals with a digital computer to effect automatic control of all required information functions -- i.e., collection, storage, access, retrieval, analysis, and display.

C. To intercept transmitted data with a digital computer and process it according to commands from the interrogating (transmitting) station.

D. To transmit the computer derived products to the interrogating station.

E. To edit, file (sort/merge) and store the data from the transmitting station. These data would be available for future access, retrieval and analysis, upon command from an interrogating station.

THE INFORMATION UNIT

Medical knowledge (including patient data) forms an exceedingly voluminous and complex body of information. One of the principal reasons for this condition is due to the fact that almost the entire body of medical knowledge is documented in natural (English, and other) language. Currently, number (quantitative) language makes up a comparatively small portion of the total file of medical knowledge. When information is stated in terms containing a set number of digits expressing absolute value, it is a very simple and straightforward matter to represent a condition or set of conditions. However, selecting word content and arranging word order to achieve specific meaning for complex biological processes is exceedingly difficult. Working under the limitations imposed by the structure of our everyday language, it is very difficult to state what we wish to state exactly and in a minimum number of words. Under many conditions it is impossible to avoid ambiguities and over wordiness.

In view of these language characteristics, a record keeping system for processing medical knowledge in a computer must be precisely structured according to strict rules. To be compatible with the structural and functional characteristics of the digital computer, the data should be organized and structured along the lines of a measurement technique such that a rigorously defined unit serves as a basis for the compiled superstructure of facts and statements. Execution of the information processing actions of collection, storage, retrieval and analysis should be a function of the measurement scheme. Additionally, one should be able to view any complex collection of information statements as an orderly series of simple data units. The process of building or reducing any information complex should be based on the information unit.

In establishing a foundation to serve as a basis for computerized record keeping it was necessary to formalize the structure of the data bases so as to provide a standardized format. The key structures of the data bases have been designated as the "item," the "message," and the "master file." However, it is the "item" which is the basic unit of the entire stystem. The item serves as a basis for standardizing all information processing carried on within the system.

In this information system the individual item of information represents the "information unit" of the system. An item of information is defined as a specific fact which can mean one and only one thing. Figures IV and V illustrate how the individual item is formated to serve as the structural unit of the data bases. Figure IV illustrates the procedure for quantitative data and figure V illustrates

the process for qualitative data. For example, the single fact captured by each item in clinical case "A," Figure V, may be stated as follows: "Cough is present"; "Sputum is produced"; Dyspnea is present"; Chest pain is present"; "Diffuse inspiratory-expiratory wheezes are auscultated"; "The forced vital capacity (FVC) is 3600 cc."; etc. No further qualification of these conditions is permitted. The simple statement of a single fact is all that is permitted. To amplify these conditions would require the gathering of additional items of information.

The items of information (FVC, MVV, MEI, etc.) presented in Figure IV in numeric values can be translated into "0" or "1" values through any one of several techniques. One of the most popular is to partition the possible ranges of the measurements into non-overlapping intervals (Figure VI). Class intervals should be developed with great care, and the statistical rules governing their construction should be followed (38, 39). Particular attention must be given to the rule that items falling into a class should be somewhat evenly distributed throughout the interval so that the mean of the items will not be greatly different from the class mark or interval midpoint. The attributes of the FVC would be partitioned into the class intervals illustrated in Figure VI. The intervals are developed from a frequency distribution analysis of the absolute numeric values illustrated in Figure IV. The other spirometry measurements could be partitioned into non-overlapping class intervals in a similar manner.

These intervals can be translated into the clinical profiling categories of "normal function," "minimal abnormality," "moderate abnormality," "severe abnormality," and "far advanced abnormality," as one progresses from high to lower performance measurements. The items of quantitative data for the clinical cases "A," "B," and "C" can then be expressed as an ordered series of "0's" and "1's" (Figure VI) in a manner identical to the items of qualitative data (Figure V). Through this technique one attains a formalized procedure for simultaneously handling qualitative and quantitative data on a common structural basis which is ideally suited to manipulation through computational processing. The computational processing can be either the traditional arithmetic operations or the propositional calculus of the modern symbolic logic. The utilization of this approach is basic to the Classification Technique for profiling disability in patients with pulmonary diseases. The Classification Technique is such a lengthy and involved computer process that it will be presented as a separate laboratory report.

The individual items of information represent the raw materials of the system and as such are regarded as a principal resource

and source of energy to the entire complex. It is convenient to view the items of information as electrical impulses or signals, and project their conversion to equivalent digital representation for processing the computer. Information items from any source can be translated into electrical signals which can be transmitted directly to the computer or stored off-line in a retention device for future input to the computer under programmed controls. In turn, computer generated signals can "feedback" to the sources of information specified orders and recommendations governing future transmissions.

The documentation of medical knowledge in natural language makes it necessary to author our documentation statements in the free style of English prose. Such statements then serve as a basis for formulating our problems, arguments and solutions. There are very few opportunities to formulate these matters in quantitative terms for computational processing through the methods of traditional mathematics. Arguments formulated in natural language, as contrasted to symbolic notation, are exceedingly difficult to evaluate because of the vague and equivocal nature of the words used, the misleading idioms they may contain, and their possibly confusing metaphorical style. Even when these problems are solved, the difficulty of determining the validity or invalidity of the argument remains.

To minimize, and hopefully to avoid in some instances, these problems, it was considered reasonable to establish within the project an artificial symbolic language as an analysis tool. The approach utilized in this project to achieve a system of symbolic logic follows the concepts, methods, and notations initiated by George Boole and ultimately extended by Whitehead and Russell (30, 31). Their work has enlarged the scope of formal logic and introduced into the process the methods and something like the symbolic notation of modern algebra. These innovations provide the basis for a computational procedure for handling information stated in non-numeric terms. Additionally, the modern symbolic logic provides a means for handling simultaneously, as terms in an expression, quantitative and qualitative information. The greater extent to which modern logic has developed its own symbolic language has made it immeasurably more powerful a tool for analysis. The writings of Ledly, et al, have also been instrumental in establishing a capability to incorporate the modern symbolic logic into the information system (32, 33).

The information unit is ideally suited to serve as a basis for implementing the notations and methods of the modern symbolic logic and its propositional calculus. Additionally, the information unit is ideally suited to function as a "simple statement" within the context of the symbolic logic (Figure VII). A simple statement

is defined as one which does not contain any other statement as a component. Every fact or simple statement is regarded to have "truth value," namely, it is either "true" or "false" for any object or condition. The condition of truth value can also be stated as being "positive" or "negative" "present" or "absent" and "yes" or "no." The state of being "positive," "present" or "yes" is represented by a "1," and the state of being "negative," "absent" or "no" is represented by a "0."

The logical basis for this assumed binary (0,1) state of information is a direct application of the "all-or-none" law for the character of nervous activity, neural events and electrical circuits (34, 35, 36). The assumed "all-or-none" law for these activities is sufficient to form a basis for a binary state of the information and to represent the activity as a proposition. Within this information system the principal requirement for a successful application of this notion is to precisely define any information item to represent one and only one fact which is strictly limited to existing in a single state of "yes" or "no."

Figures V-VII illustrate the basis that is established through the information unit to implement a logical calculus as established by George Boole, and a calculus of propositions as established by Whitehead and Russell (30-31). Although the illustrations presented in these figures are simple and straightforward, the system is entirely capable of developing most any degree of complexity in statement compounding and argument formulation that is required. George Boole established a calculus in which the only values of the variables are "0" and "1." Claude Shannon was the first to apply this logical calculus to nets of electrical relays (34). However, Shannon was interested only in the state of circuits being open or closed rather than in transients opening and closing circuits. McCulloch and Pitts extended the ideas and methods of these authors by taking into consideration the factor of "time," wherein the unit of time is synaptic delay or closing time (35-36). They made a complete calculus for these signals by taking the calculus of propositions of Whitehead and Russell from the "Principia Mathematica" and subscripting the symbol for the signal of a given relay by the time of that signal measured in synaptic delays from any arbitrary beginning (36). In their opinion this calculus was much simpler than ordinary arithmetic and enormously more powerful (36).

The development of the information unit method to serve as a basis for the computerized biomedical information system is in its preliminary phase. At the present time the method has been almost entirely restricted to the subspecialty of pulmonary diseases. It does appear to hold promise as a general tool for approaching

computer applications in the entire medical field. As the scope and complexity of this information system increase the information unit method must be further developed to meet the needs and fulfill the obligations of the system. If the method cannot function in a universal capacity it must be limited to its areas of proven success, or replaced by a more general approach.

MESSAGE FORMATING AND THE MASTER FILE

Information items are passed into a collection system where they are combined into cohesive complexes or messages. The collection system is comprised of the computer programming which perform the data handling required to format individual items into messages. This action takes place in the computer high speed memory. A "message" is defined as a compound statement containing more than one fact. It may also be viewed as a polynomial expression with multiple terms or factors. The items of information comprise the "terms" or "factors" of the series. Each item for any message is selected on the basis of medical criteria. For example:

1. The "Spirometry" message is composed of information items which are selected on the basis of their proven role in evaluating the state of pulmonary ventilations.

2. The "Arterial Blood Gas" message is composed of information items which are selected on the basis of their proven role in evaluating the oxygen-carbon dioxide-pH state of the arterial blood at rest and during graded exercise.

3. The "Respiratory Disease Questionnaire" message is composed of information items which are selected on the basis of their proven role in documenting past and present respiratory illnesses and current respiratory symptoms. These items are of proven value to the physician evaluating pulmonary diseases.

The message represents a specific subject matter category, and it is intended to provide a factual description of a given body structure, physiological or biochemical process, disease process, treatment regimen, medical procedure, etc. In selecting the items it is necessary to anticipate the information which is required for control or problem solving within a specific subject category.

Items of information are generated at multiple sources. Items must be collected, ordered into series and filed in preparation for their use by subject matter specialists. To structure a complete and logically cohesive message describing a specific subject category, such as external pulmonary ventilations, the essential items of

information must be collected from various data sources. The required items may come from such diverse sources as the doctor's office, the nurse's station, the ward, and the pulmonary function laboratory. The collection system effects their union into a single descriptive statement of external pulmonary ventilation.

A system intended to service multiple users must have the flexibility to respond to the various preferences of individual users. The inclusive series of items considered to be a full and accurate description of a subject by one user may not be accepted by others. Through the use of a program logic system designated the "Population System" any, or all, items of a message can be joined to any, or all, items of other messages to produce a new message. The "Population System" is especially designed to search any number of data files, access information items on the basis of imposed criteria, format a new message from existing messages and compile the required number of messages into a new master file. To relate the master file information to mathematical and statistical analysis, a question must be asked. For instance, a user might ask:

1. What is the expected forced vital capacity (FVC) in a normal male?

2. What is the expected FVC in a normal male non-smoker?

3. How does the mid-expiratory time (MET) correlate with the functional residual capacity (FRC) in the normal male non-smoker?

To answer these types of questions, we need computer programming that can produce means and standard deviations, regressions and correlations from files of data. However, we also need computer programming that will select and format the data for the statistical programming to operate on. This is the function of the Population System.

Within the computer high speed memory messages are identified, sorted and compiled into specific subject matter "master files." A master file can be defined as a data block composed of a collection of messages of a specified category. A master file is characterized by the following properties:

1. Each master file contains cohesive units of information, the message, such as spirometry, smoking history, lung volumes, arterial blood gases, respiratory disease questionnaires, surgical procedures, etc.

2. Any individual may be represented any number of times in any file.

3. An individual is assigned one code for all files.

4. In a master file the messages are of the same length or variable length and are in ascending code sequence.

5. There is unlimited expansibility to a master file. The master files are written from computer memory to external storage on magnetic tape, paper tape, 80 column cards or magnetic discs for random access. The storage, analysis, retrieval, display and quality control editing functions involving the master files are processed under automatic programmed controls. The exclusive computer code assigned to each individual offers easy, fast programmed access to the data available on any individual. At the present time the data base in the information system is composed of 58 discrete master files (Table I). Since there is unlimited expansibility to the data base this total is subject to continuous change.

The raw data required to establish the Spirometry master file is collected in the Pulmonary Function Clinic, Fitzsimons General Hospital. Technicians in the Pulmonary Function Clinic administer the test to the patients, carry out the required measurements on the respiratory tracing and record the numerical values on the Spirometry Raw Data Input Format (Figure X). The items of raw data are used to calculate the desired physiological measurements of pulmonary ventilation which include:

1. Forced Vital Capacity (FVC mm.).

2. One-second Forced Expiratory Volume (FEV_1 mm.).

3. Mid-expiratory Time (MET mm.).

4. Maximum-mid Expiratory Flow Rate (MMEF mm.).

5. Maximum Expiratory Flow Rate (MEFR mm.).

6. Maximum Ventilation Volume (MVV Liters/min. mm.).

7. Maximum Ventilation Volume, Respiratory Frequency (MVV FREQ.).

Staff personnel in the Pulmonary Function Clinic obtain and record clinical items of information (Figure X) which are used to perform clinical correlation studies and derive regression equations

for computing predicted values for observed values of pulmonary ventilation. The clinical items include:

1. Standing height (Ht. in.).

2. Weight (Wt. lbs.).

3. Age.

4. Administration of the bronchodilator, Isuprel, 1:200 solution, under intermittent positive pressure breathing (IPPB) (Isuprel).

5. Functional classification (FC).

6. Status of the individual tested as to being a "normal" subject or a patient (Norm Study).

7. Quality of the respiratory tracing (Tr Qual).

8. Number of the Spirometry test (Spiro No.).

9. Code for the physician interpreting the study (Tag).

10. Code for the technician who performed the test (Tech).

The Spirometry Raw Data Input Format is transferred to the Computer Division, USAMRNL, daily at 1500 hours. Staff personnel in the Computer Division write the data from the input format to punched paper tape with a Friden programmatic flexowriter-verifier unit. In addition to a machine processed verification routine further quality control of the data is obtained by a read back of the flexowriter print-out to an individual who reads from the original data input format.

The verified punched paper tape is placed in the paper tape reader and, under automatic programmed controls, the following operations carried out:

1. The information is read to the computer high speed memory.

2. The required computer programs are called into high speed memory from magnetic tapes.

3. The physiological measurements of ventilation (observed values) are calculated from the raw data. The values are checked for "reasonableness" against imposed criteria and if a value fails the test and error halt is generated.

4. The predicted values for the physiological measurements of ventilation are calculated from regression equations. The regressions are derived from a population of 875 normal subjects. These equations are periodically derived anew as additional cases are added to the population set and some cases are deleted because they no longer qualify as normal subjects because of changes in health status or revision in criteria for "normal."

5. The numerical values of the ratios of observed values/predicted values are calculated.

6. The items of information are formated into Spirometry master file format (Figure IV).

7. Each message is ordered into its proper place (code sequence) in the master file through the special purpose Sort/Merge routine. This routine was designed and developed in the Computer Division for meeting the needs of quality control and updating of patient files.

8. Clinical reports (Figures XI, XII) which are to be distributed to the hospital physicians are formated and processed on the high speed printer.

The sequence of messages in Figure IV illustrates the format of the data as it exists in the computer high speed memory during operational processing. It is convenient to view this collection of messages as an ordered series of rows and columns (a matrix) wherein the clinical cases are arranged in rows and the information items are arranged in columns.

The spirometry message contains 23 items of information. It is a "fixed" length message as it always contains exactly 67 alpha-numeric characters. Each item states one and only one fact and can be viewed as a simple statement with truth value. The message contains all the information obtained from a single spirometry test.

The master file is open ended and can contain as many spirometry messages as the user requires. A specific message can appear only once in the file. A specific case can be represented by as many different spirometry messages as the user requires. Messages can be added or deleted under automatic programmed controls upon demand by the user.

The selection of items of information to be compiled into messages is based upon medical criteria. The items of information

in this file have been selected on the basis of their proven role in evaluating a variety of diagnosis-treatment problems involving diseases of the chest.

Some of the roles are:

1. Providing a means for correlating subjective complaints of dyspnea with numerical values of pulmonary ventilation.

2. Uncovering the early presence of a disease process by demonstrating the presence of impaired pulmonary function at a time when information from the history, physical examination, chest x-ray, and other studies remain within normal limits. This has proven to be particularly true of early airway disease which is manifested by airflow obstruction with reduction in the FEV_1/FVC ratio, the maximum-mid expiratory flow rate and in prolongation of the mid-expiratory time.

3. Providing assistance in evaluating "poor risk" patients who are candidates for thoracotomy and pulmonary resection.

4. Providing numerical standards for evaluating pre and post surgical (pulmonary resection) health status.

5. Providing objective, standardized criteria for profiling performance disability in patients with cardio-pulmonary disease, certain neuro-musculo-skeletal diseases, and anxiety states.

6. Providing objective criteria (numerical values) for evaluating response in pulmonary ventilation to therapy.

7. Providing standardized criteria for following the natural history of pulmonary disease.

These procedures for establishing, maintaining, and utilizing the Spirometry master file are generally common to all the other files (Table I) which currently comprise the data bases for the information system. A patient master file has been established for the convenience of assembling in one series all medical information for any patient. Although the Master File appears as the super structure of the information system, it is the "information item" which provides a basic unit for executing all processes (storage, retrieval, analysis, display) in a logical manner compatible with the hardware structure and function.

COMPUTER PROGRAM LIBRARY

The single most demanding requirement continuously encountered throughout the project is the development of computer programming.

The design, authoring and debugging of computer programming is a difficult and time-consuming effort. Depending upon the scope and complexity of the process being programmed, days, weeks, months, or years may be required to establish an operational program. A routine to produce a print-out of high speed memory in RCA 301 program format was completed within several days. A routine, "The Population System," to access specified master files, interrogate them for specified information items, retrieve the data, establish a new master file and format the data for statistical analysis has been under construction for several years. The development of such a program system is characterized by successive achievements in expanding its capabilities and improving its generality. However, it never does attain a "perfected" or "completed" status. If such a status were achieved the program would, at long last, attain a status of complete generality.

The primary objective of the program library is to permit the user to delegate the physician's functions to the automata through effective, accurate simulation of medical reasoning foundations. The approach utilized to achieve this objective is based upon two basic procedures:

1. Design a logic structure which serves as a guide for medical problem solving.

2. Author a library of computer programs which permits the rules and policies of the logic structure to be executed under automatic controls.

For purposes of this project the logic structure for medical problem solving is viewed as a multistage process requiring a sequence of decisions over space and time. The notion assumes that problem solving is not a one-shot affair, but rather involves many interconnected observations, thoughts and actions. The general strategy of the multistage procedure is as follows. The individual views a situation and examines it in terms of its needs and obligations. Once the problem is defined one now chooses one from many available courses of actions. The initial course of action is based upon an interaction between "information" and "reasoning foundations." At subsequent times further examinations are made to evaluate the effectiveness of preceding decisions, and to obtain the information and formulate the reasoning foundations required for the next decision. This complex ("The Course of Action") of gathering information, formulating reasoning foundations, and decision making is repeated until a solution to the problem has been established. In some cases this complex must be repeated indefinitely because there is no finite solution to the problem. In such cases a need exists to continuously update some previous decision. This method of

sequential decision processes is also common to medical research and many other control and problem solving processes described elsewhere (37).

Guided by this logic structure for medical problem solving the library of computer programs has been designed and developed to assist the physician or research investigator as he engages in the process of sequential decision making. The library is in fact a series of systematic techniques for handling information according to the dictates of the user's interrogations. The user's interrogation is a formal expression of concern about some aspect of the problem and determines what information is required and how it should be used effectively. The program library is designed to respond to the needs of the user through this process of interrogation or "framing the question." For example, the physician engaged in treating a patient with pulmonary tuberculosis may, in the process of evaluating the patient's capability to undergo major thoracic surgery, raise the question, "What is the expected change in pulmonary function following a total right upper lobectomy in this 55-year-old male with bilateral upper lobe tuberculosis?"

Figure VIII illustrates the scheme for effecting an interaction between medical problem solving and the computer program library. The nature of the interaction determines the content of the program library. The program library's simulation of the multistage sequential process of decision making is approached by establishing a closed-loop system which can iterate any finite number of times required by the user. Additionally, the user can "interrupt" the cycle at any point in the process, access the desired process station through the program library and "lock" the action at that process station through any desired period of time. This is actually carried out by executing a dump of the computer high speed memory to tape or discs for storage until such time as the user is ready to continue with the process. However, there is only one pathway for legitimate exit from the closed-loop, namely, by way of the decision station identified as "Problem Solved?".

Each of the major processing stations within the closed-loop (Figure VIII) are supported by a detailed battery of programs (principal routines and subroutines) employing systematic techniques to satisfy the needs and fulfill the obligations of that station. For purposes of this discussion the terms "program" and "routine" are used interchangeably and are defined as a set of coded instructions arranged in proper order to direct the computer to perform a desired operation or series of operations. A subroutine is defined as a set of instructions used to carry out a well-defined mathematical or logical operation and thereby functions as a subunit of a routine.

Table II presents an inventory of the library of computer programs classified according to the principal processing stations represented in Figure VIII. The programs listed in Table II are written in four programming languages, namely, absolute machine, Fortran II, Assembly System and University of Miami Algebraic Compiler (UMAC). The number of instructions per program is determined by two components, namely, the principal program logic and its supportive (usually floatable) subroutines. For example a program to develop a Spirometry Master File containing observed and predicted values (Table II, Compile Master Files, Program No. 1) has 750 instructions in the principal program logic, and 227 instructions (floatable) in the floating decimal arithmetic --8 digit words-- (FDA-8) subroutine.

The date column in Table II gives the date of the most recent rewrite or debugging of the program. For example, the program "Classification No. 1)SRM16)" was first run in 1965, but the most recent rewrite and debugging is listed as 18 April 1968. It is most likely that this routine will undergo further modifications and rewriting in the future. This will establish a new date citation for this program in Table II. Many of the programs have gone through several stages of rewriting.

The requirement to design and author new programs and to modify and rewrite existing programs is a never ending one. This requirement correlates most strongly with the well-known perversity of the human organism when involved in a disease process. The characteristic of meandering or straying of physiological, biochemical and psychological processes throughout the course of illness is a principal consideration in authoring computer programs capable of meeting the needs and obligations of the user. The user usually develops a projected pathway for the natural course of the illness, and then formulates his needs and obligations accordingly. Figure IX-a, illustrates the notion of a physician establishing a projected pathway of the natural course of a disease process. The pathway is represented as a continuous change in the illness over time in moving from the initial health state (Hi) to an ultimate normal health state (Hf). Since medical practice is not an absolute science but rather relative portions and degrees of art and science, the pathway "Hi-Hf" is not a quantified, calculated trajectory. As the well-known perversity of the human organism begins to operate any number of malfunctions (small to large) can accumulate and produce serious deviations from the projected trajectory. Figure IX-b, illustrates a state of meandering from the projected pathway "Hi-Hf" along a course "p^1 -Ha." Although many of the possible deviations can be assumed in advance, it is just not possible to anticipate the almost infinite number and variety (combinations) that can occur in all patients. The physician must continuously respond to this zig-zag course through the gathering of additional information and formulation of alternative reasoning foundations.

In view of these phenomena, the library of computer programs must provide the techniques for dealing with the complications in information processing which result from this biological perversity. In addition to handling the projected pathway, the program library must contain as many corrective procedures as are required to handle the deviations which occur around the projected clinical trajectory of a disease process. As the number and variety of clinical cases in the patient files increases, the requirement for correction techniques becomes greater and greater. It is unrealistic to plan a library of computer programs for the domain of biomedical applications that does not recognize this basic biologic phenomenon in human beings stricken with illness. Recognition of this phenomenon will immediately dispel any illusion one might have regarding any "easy" and "quick" development of a computer program library for biomedical applications. Considering the present state of the art and the austere scarcity of usable products for patient care, it is most realistic to assume that the medical profession faces a monumental task to develop effective computer programs for medical applications. Additionally, knowledge and experience about this phenomenon should convincingly dispel any illusion that one can obtain from another institution a complete computer program library for medical applications in patient care.

Looking ahead into the future it is rational to visualize how these preparations can serve as a basis for a system of "dynamic programming" which would attempt to formulate rules and policies for determining the best decisions to make in order to solve the problem (37). As herein used, "dynamic programming" would be defined as a sequence of step-by-step operations which are to be performed by the computer to develop rules and policies which determine what decisions should be made in terms of the current position of the disease process. Under this concept of "dynamic programming" of disease course control, the intent of the "correcting" influence of the program would be to project a revised trajectory to a normal health status extending from the point to which the patient had strayed from the original clinical trajectory to the desired state of normal health. Figure IX-c, illustrates this notion in selecting a revised pathway "Ha-Hf" in preference to returning to the original pathway "Hi-Hf" in order to attain the ultimate health state, Hf. The major advantage of this control concept is that it provides unlimited flexibility in dealing with all eventualities. Regardless of the current position of the health state a policy is developed which determines the optimum way to attain the ultimate goal. Under this concept we are not bound by preconceived notions which would always require that we return to the original projected pathway (Figure IX, "Hi-Hf"). The concept of "dynamic programming" advocates

the basic notion of learning from experience in finding the most desirable path to achieve the ultimate goal.

Each computer program listed in Table II is on file in the Computer Division and is documented according to the following outline:

1. General flowchart.
2. General description.
3. Input preparation.
4. Operating requirement.
5. Normal operating procedure.
6. Normal operating conditions.
7. Abnormal operating conditions.
8. Sample problem.

Despite differences in hardware and program coding, it is felt that the functional characteristics of these routines can be made available to others through the transfer of logic flow charts and discussions between appropriate personnel. In matters concerning "systems programming" the transfer of concepts and the logic of an approach to the problem is certainly equal to that of machine coding. Systems programming is continuously in need of alteration, updating and rewriting. Sooner or later every user faces up to the fact that he must author computer programming to keep his system viable.

SUMMARY

A computer based biomedical information system is being developed to provide automatic data processing services to clinical physicians and research investigators. The development of a general system for clinical medicine and laboratory research is approached through an exercise utilizing a small scale prototype model "The Classification of Disability in Pulmonary Diseases."

The system has achieved a state of readiness to handle the information processing functions of data collection, storage, retrieval, analysis and display under relative degrees of automatic control. The principal requirement to expand and perfect the capabilities

for automatic data processing is the authoring of computer programming. Currently the computer program library contains 67,098 instructions, and new programming is being produced at an average rate of 64 instructions/day or approximately 16,600 instructions/year. As the computer program library expands the system achieves an improved state of "generality" which serves to increase the number and scope of medical applications.

To realize a precise simulation of the information processing activities involved in clinical medicine and laboratory research, the system must achieve a state of readiness to serve physicians and investigators through real-time, on-line operations. The research program is progressing towards this realization and it is anticipated that the system will achieve a state of readiness to implement "limited" real-time, on-line operations in approximately one year.

TABLE I MASTER FILES

NO	TITLE
1	Code/Name - Code Sequence
2	Code/Name - Name Sequence
3	Symptomatic Complaints - Respiratory Diseases
4	Urinalysis
5	Hematology
6	Spirometry Predictions
7	Ear Lobe Oximetry
8	Arterial Blood Gases - 4 step
9	Allergy Skin Tests - Inhalants
10	Allergy Skin Tests - Mixes
11	Respiratory Disease Questionnaire
12	Exercise Step Test
13	Diffusing Capacity - Steady State
14	Diffusing Capacity - Single Breath
15	Standard Form - 88
16	Standard Form - 88 - Supplementary
17	Standard Form - 89
18	Standard Form - 89 - Physician's Pertinent History
19	Smoking History
20	Spirometry
21	Spirometry Predictions & Ratios Observed/Predicted Values
22	Military Register

TABLE I MASTER FILES

NO	TITLE
23	Lung Volumes
24	Lung Volumes - Predicted Values
25	ECG Interpretations
26	Respiratory History
27	Physical Examination - Pulmonary
28	Arterial Blood Gases - Single Blood
29	Fluoroscopy
30	Surgical Procedures
31	Diagnoses - General
32	Diagnoses - TBC
33	Diagnoses - Detailed
34	Medical Terms - Term Sequence
35	Medical Terms - Code Sequence
36	Microbiology Data Format - No. 1
37	Microbiology Data Format - No. 2
38	Microbiology Data Format - No. 3
39	TBC Section Conference Data
40	TBC Current Therapy Data
41	TBC Initial Therapy Data
42	TBC Admission & Contact Data
43	TBC Respiratory History
44	TBC Chest X-ray
45	TBC Chest X-ray

TABLE I MASTER FILES

NO	TITLE
46	TBC Diagnosis & Disposition
47	TBC Discharge Data
48	TBC Follow-up Therapy Report
49	TBC Laboratory Data
50	TBC Thoracic Surgery Data
51	OB-GYN Questionnaire - No. 1
52	OB-GYN Questionnaire - No. 2
53	OB-GYN Questionnaire - No. 3
54	OB-GYN Questionnaire - No. 4
55	Allergy Clinic - Asthma Study
56	Tumor Registry, FGH - Epidemiological Data
57	Tumor Registry, FGH - Smoking History
58	Tumor Registry, FGH - Diagnosis
59	Tumor Registry, FGH - Surgical Procedures
60	Tumor Registry, FGH - Cobalt 60 Irradiation
61	Tumor Registry, FGH - External Radiotherapy
62	Tumor Registry, FGH - Radium Therapy
63	Tumor Registry, FGH - Strontium 90 Irradiation
64	Tumor Registry, FGH - Hematology
65	Tumor Registry, FGH - Special Chemistry
66	Tumor Registry, FGH - Serum Protein
67	Classification No. 1
68	Classification No. 1 - Attribute Coefficients

TABLE II COMPUTER PROGRAM LIBRARY

MESSAGE FORMATS

NO	TITLE	INSTRUCTIONS	DATE
1	Change Coded Term In Any Message In Patient Master File	484	8 May '67
2	Eliminate Messages From Any Master File On Imposed Criteria	100	15 Mar '66
3	Execute Message Selection From Any Master File On Imposed Criteria	256	22 Apr '67
4	Remove Duplicate Messages From Patient Master File	145	26 Feb '68
5	Edit Messages In Lung Volume Master File	100	24 Jan '67
6	Edit Messages In Standard Form -88 Master File	91	20 Jan '67
7	Edit Messages In Respiratory Disease Questionnaire Master File	57	23 Jan '67
8	Edit Messages In Smoking History	58	23 Jan '67
9	Edit Messages In Respiratory History	55	16 Feb '67
10	Edit Messages In Physical Examination, Pulmonary Diseases	69	15 Feb '67
11	Edit Messages In Standard Form -89 Master File	74	10 Feb '67
12	Execute Item Correction For Any Message	224	14 Jul '66
13	Verify Message Length on Identification Numbers In Patient Master File	104	12 May '67
14	Retrieve Messages by Date	53	2 Oct '67
15	Check Message Order and Duplicate Codes	48	22 Jun '67
16	Sort Messages Within A Master File By EB's	84	7 Aug '67
17	Execute Message Selection From Code-Diagnosis Master File	67	4 Oct '66

TABLE II COMPUTER PROGRAM LIBRARY

MESSAGE FORMATS

NO	TITLE	INSTRUCTIONS	DATE
18	Specific Message Selection From Individual Master File To Individual Master File	78	31 Aug '67
19	Message Set Selection	100	21 Aug '67
20	Code/Message Selection From Master File Format To Master File Format	102	18 Aug '66

COMPILER SYSTEM

NO	TITLE	INSTRUCTIONS	DATE
1	Assembly Language Translator	350	Mar '68
2	Fortran II	16,200	Apr '68
3	University of Miami Algebraic Compiler	13,250	Nov '67

GENERAL SERVICE ROUTINES

NO	TITLE	INSTRUCTIONS	DATE
1	Program Library Tape	499	Feb '68
2	Program Library Tape, Update	15	20 Jan '68
3	High Speed Memory Dump Routine	14	21 Feb '68
4	Duplicate Magnetic Tapes	45	15 Jun '66
5	Geopolitical Area Codes	29	30 Jun '67
6	Predicted Values Tables	109	28 Oct '66
7	Format Data On Punched Paper Tape	40	27 Jan '67

TABLE II COMPUTER PROGRAM LIBRARY

DATA DISPLAY

NO	TITLE	INSTRUCTIONS	DATE
1	Microbiology Data Sheet No. 1 Print-Out	456	15 Apr '68
2	Data Reduction Graphics	350	13 May '68
3	Classification No. 1, Print-Out	205	28 May '68
4	Spirometry Master File Print-Out	68	4 Aug '67
5	Spirometry Predictions Print-Out	140	15 Nov '67
6	Physical Examination, Specified, Print-Out	119	21 Jul '66
7	Standard Form -88, Print-Out	122	2 Dec '66
8	TBC Admission & Control Data, Print-Out	116	14 Dec '67
9	Arterial Blood Gases, Print-Out	166	17 Nov '67
10	Respiratory Disease Questionnaire, Print-Out	116	1 Jul '66
11	Smoking History, Print-Out	107	30 Jun '66
12	X-Ray Interpretations, Print-Out	44	3 Jul '65
13	ECG Interpretations, Print-Out	42	8 Jul '65
14	Any Specified Patient Or Master File Print-Out In Special Format	3301	20 Dec '67
15	TBC Current Therapy Data, Print-Out, Part I	167	9 Feb '68
16	TBC Current Therapy Data, Print-Out, Part II	122	15 Feb '68
17	All Punch Conversion Listing	176	8 Feb '68
18	Clinical Information Report	836	20 Oct '66
19	Print A Patient's Record To High Speed Printer	183	26 Oct '67
20	Individual Master File Print-Out	136	25 Nov '66
21	Medical Terms, Print-Out	92	15 Aug '65
22	Patient Master File Print-Out, Master File Format	23	22 Jun '66

TABLE II COMPUTER PROGRAM LIBRARY

DATA DISPLAY

NO	TITLE	INSTRUCTIONS	DATE
23	Print Ending Tape Labels, Beginning Tape Labels, First & Last Messages	27	16 Jun '66
24	Spirometry Prediction, Print-Out, Code & Name	292	17 Jan '68
25	Print-Out High Speed Memory In Program Format	40	30 Dec '68
26	Individual Master File Print-Out by EI'S	36	7 Jun '66
27	Print Another Floatable Program	29	16 Jun '67
28	Print-Out Messages From Individual Master Files For Supplied Codes	31	23 Aug '67
29	Produce and Print-Out A Medical Summary For Any Specified Clinical Case	2232	20 Aug '67
30	Code/Name Master File Print-Out	69	14 Jun '66
31	Pathology Division Master File Print-Out	74	8 Nov '66
32	Medical Terms Print-Out	27	11 Oct '66
33	OB-GYN Questionnaires No. 1, 2, 3, 4 Print-Outs	552	25 Mar '67
34	Lung Volume Master File Print-Out	103	1 Jul '66
35	Spirometry Master File Print-Out, Master File Format	68	22 Jul '67
36	Spirometry Predictions Print-Out, Master File Format	140	15 Nov '67
37	Physical Examination Master File Print-Out	119	21 Jul '66
38	Standard Form -88, Master File Print-Out	122	2 Dec '66
39	TBC Admission & Contact Data Print-Out	116	21 Dec '67
40	Respiratory Disease Questionnaire, Master File Print Out	110	1 Jul '66

TABLE II COMPUTER PROGRAM LIBRARY

MANIPULATE DATA

NO	TITLE	INSTRUCTIONS	DATE
1	Equation Solver	160	1 May '63
2	Arithmetic Mean	85	1 May '63
3	Standard Deviation	110	1 May '63
4	Simple (Personian) Correlation	140	25 Jul '65
5	Multiple (Personian) Correlation	185	25 Jul '65
6	Multiple Regression	190	25 Jul '65
7	Analysis of Variance	210	25 Jul '65
8	Analysis of Covariance	280	25 Jul '65
9	Matrix Add (Subtract)	39	1 May '63
10	Matrix Scalar Multiplication	28	1 May '63
11	Matrix Multiplication	90	1 May '63
12	Matrix Transpose	87	1 May '63
13	Matrix Equation Solver	150	25 Jul '65
14	Matrix Inversion	195	25 Jul '65
15	Floating Decimal Arithmetic -8	227	1 May '63
16	Floating Decimal Arithmetic -13	247	1 May '63
17	Floating Decimal Arithmetic -18	265	1 May '63
18	Scientific Data Edit	361	25 Jul '65
19	Matrix Transcribe (Fortran II Statements)	150	1 Oct '67
20	Matrix Determinant	150	25 Jul '65
21	Spirometry Predictions for Classification No. 1 (SRM11)	514	28 Jan '68
22	Spirometry Predictions and Ratios (SRM12)	750	4 Mar '68
23	Classes for Spirometry Variables (SRM13)	505	16 Feb '68
24	Classification No. 1 (SRM16)	369	18 Apr '68
25	Frequency Distribution, Vital Capacity	190	27 Apr '67
26	Domain of Numbers, Classification No. 1a	154	4 Feb '66

TABLE II COMPUTER PROGRAM LIBRARY

MANIPULATE DATA

NO	TITLE	INSTRUCTIONS	DATE
27	Frequency Distribution, Age	246	21 Apr '65
28	Frequency Distribution, Height	271	3 Sep '65
29	Compute Age & Height Categories	270	15 Oct '66
30	Classification No. SP-1	520	21 Nov '67
31	Frequency Distribution, Maximum Ventilation Volume (MVV)	176	21 Oct '67
32	Convert Classification SP-1 to Binary State (SRM08)	330	23 Oct '67
33	Compute Total Probabilities on Respiratory Disease Questionnaire (RDQ)	105	3 Apr '63
34	Frequency Distribution, Illnesses and Symptoms on RDQ	359	25 Apr '63
35	Gross Nutrient Balance	325	11 Jun '65
36	Format Statistical Analysis System (SAS)	283	5 Mar '68
37	Convert Calendar Dates to Julian Dates	44	4 Apr '68
38	Establish Classification SP-1 Attributes	430	12 Apr '67
39	Combine Attributes for Classification No. 1 (SRM14)	260	16 Mar '68
40	Total Body Water	192	12 Jan '68
41	Analog-Digital Conversion Tabulation	53	12 Oct '67
42	Energy Expenditure	875	16 Sep '65
43	Translation Table, 501 Language to 301	343	20 Dec '62
44	Convert "Day of Year" to Month & Year	23	27 Apr '65
45	Convert Julian Date to Calendar Date	42	27 Apr '65
46	Calculate Pack/Yrs of Smoking	60	31 Mar '67
47	Recalculate Spirometry Per-cent Predictions	101	15 Nov '64
48	Statistical Analysis and Classification of Spirometry	598	15 Oct '67

TABLE II COMPUTER PROGRAM LIBRARY

MANIPULATE DATA

NO	TITLE	INSTRUCTIONS	DATE
49	Cardiac Computations (Fortran II Statements)	515	21 Oct '66
50	Floating Decimal Arithmetic -8, Fixed to Floating Decimal Data Conversion	99	15 Apr '65
51	Floating Decimal Arithmetic -8, Floating To Fixed Decimal Data Conversion	72	4 Apr '65
52	Frequency Distribution & Histogram	364	2 Mar '67
	MASTER FILES		
1	Sort And Merge Messages Into The Patient Master File	188	22 Sep '66
2	Expand or Contract Size in Patient Master File	95	6 Nov '67
3	Establish Individual Master File With Patient's Name	170	29 Jan '68
4	Update Patient Registry With Place & Date of Birth, and Social Security No.	142	11 Apr '68
5	Update Individual Master File	51	22 Jul '66
6	Individual Master File Generation	22	16 Nov '66
7	Create Prime Normal File	165	20 Sep '66
8	Establish Master Files On Magnetic Tape From Magnetic Discs	42	29 Sep '66
9	Calculate Patient Master File Message Tally	80	2 Mar '67
10	Create Subfile Selecting Messages In Specific Locations	139	22 Apr '65
11	Establish A Master File On Magnetic Tape From Paper Tape	41	5 Jan '67
12	Executive Control Routine	16	15 Jan '67
13	OB-GYN Follow Up	55	4 Dec '67
14	Arterial Blood Gases Field Count	77	8 Dec '67
15	Format Patient Master File To SAS Format	174	15 May '67

TABLE II COMPUTER PROGRAM LIBRARY

MASTER FILES

NO	TITLE	INSTRUCTIONS	DATE
16	Convert Master Files From Magnetic Disc Format to Magnetic Tape Format	132	1 Jun '66
17	Prepare Master File For Statistical File	186	8 Jul '66
18	Compile Corresponding Messages From Spirometry, Lung Volumes and Spirometry Prediction Master Files	198	22 Jan '68
19	Edit Master Files	240	11 Jan '68
20	Format Master File Data From Magnetic To Punched Paper Tape For Fortran II Insert	60	18 Jan '67
21	Population System Linkage (Control)	72	22 Sep '66
22	Update Code/Diagnosis Master File	245	20 Dec '66
23	Merge Any Two Master File Tapes	74	16 Oct '65
24	Eliminate Any Master File	67	16 Oct '65
25	Multiple Tape Duplication	144	26 May '67
26	Change Master File Block Size	39	16 Jun '66
27	Master File Daily Update	3417	3 Aug '66
28	Expand Master File, Standard Form -88 and Change Any Specified Field	148	18 Mar '68
29	Abstract Magnetic Tape	52	23 Jan '68
30	Convert Files From Magnetic Tape to Paper Tape	67	26 Jan '68
31	Delete A Specified Number of Master File Blocks	138	3 May '68
32	Master File Program Library - Select Desired Program via Parameter Message	20	3 Nov '67
33	Edit Master File (SRM11)	76	11 Dec '67
34	Establish Classification Attributes in Full Blocks	286	7 Apr '68
35	Establish Classification -1a in Master File Format, Store On Magnetic Tape	334	24 Apr '68

TABLE II COMPUTER PROGRAM

MASTER FILES

NO	TITLE	INSTRUCTIONS	DATE
36	Sort/Merge For Any File or Message	2300	25 Jul '65
37	Cross-File Interrogation By Imposed Criteria To Select Special Messages	150	31 Aug '66
38	Retrieve Data From Patient Master File By File Identification Number	37	3 Apr '67
39	Compare Lung Volume File to Spirometry File For Matching Test	74	21 Jul '66
40	Create An Individual Message Master File From Patient Master File	118	25 Nov '67
TOTAL INSTRUCTIONS, PROGRAM LIBRARY (Currently the Average Production of Operational Computer Program Instructions is 64/day or 16,600/year)		67,098	

BIBLIOGRAPHY

1. Lipkin, Martin, Engle, Ralph L., Zivorykin, V. K., Ebald, R., and Sendrow, M.: Digital Computer as Aid to Differential Diagnoses. Archives of Internal Medicine, 108: 56-72, July 1961.

2. Zivorykin, V. K., and Berkley, C.: Applications of Medical Data Processing. Presented at Instrument Society of America, National Instrumentation Conference, Sep. 9, 1957, Cleveland, Ohio, Paper No. V Me-3-57.

3. Malindzak, George S. Jr., and Stacy, Ralph W.: Dynamic Behaviour of a Mathematical Analog of the Normal Human Arterial System. Am. Jr. Med. Electronics, 4 :28-34 Jan-Mar 1965.

4. Berson, Alan S., Stallmann, Friedman W., Brodus, John H., and Pipberger, Humber T.: Telephone Transmission of Electrocardiograms and On-Line Computer Diagnosis. Am. Jr. Med. Electronics, 4: 35-37 Jan-Mar 1965.

5. Berkley, Carl: Case Histories - An Untapped Medical Information Resource. Am. Jr. Med. Electronics 4: 4-5 July-Sept. 1965.

6. Tanimoto, T. and Loomis, R. G.: The Application of Computers to Clinical Medical Data. IBM Medical Symposium, June 13-17, 1960.

7. Lindberg, Donald B. B.: Electronic Retrieval of Clinical Data. Jr. Med. Education, 40: 753-758 Aug. 1965.

8. Spencer, W. A. and Vallbona, C.: Applications of Computers in Clinical Practice. JAMA 191: 121-133, 1965.

9. Gustafson, J. E. and Throckmorton, T. D.: The Computer as a Consultant. Med. Times 93: 115-119, Feb. 1965.

10. Heaf, P. J.: Automation In Medicine. Proc. Roy. Soc. Med. 57: 1148-1149, Dec. 1964.

11. Bennett, H. E. and Holland, W. W.: Towards the Development of Electronic Data Processing Systems for Medical Records. The Lancet, 289: 1176-1178, December 4, 1965.

12. Arnold, R. F., Garner, H. L., Karp, R. M., and Lawley, E. L.: Mathematical Models of Information Systems. Technical Report No. RADC -TR -66-37, April 1966.

13. Technical Meeting on Design, Implementation and Application of Information Retrieval Oriented Languages. Comm. ACM. 5: 8-47, 1962.

14. Baruch, J. J.: Progress In Programming For Processing English Language Medical Records. Ann. N. Y. Acad. Sci., 126: 795-804, July 1965.

15. Jenkins, David R.: Problems of Computer Applications in Medical Research. Transactions of the New York Academy of Sciences, 28: 439-447, January 1966.

16. Warner, Homer R.: Editorial. Computer and Biomedical Research, 1: 1-2, 4 February 1968.

17. Wilson, Ira G. and Wilson, Marthann E.: Information, Computers, and System Design. John Wiley & Sons, Inc., New York-London-Sydney. p. 3-20, 1966.

18. Memorandum 30-3, U. S. Army Adjutant General School: Introduction to Automatic Data Processing System (ADPS), February 1965.

19. Memorandum 32-1, U. S. Army Adjutant General School: Introduction to Systems Analysis, July 1965.

20. Radio Corporation of America (RCA) Manual: An Approach to the Basic Techniques of the Systems Analysis.

21. Gravenstein, J. S.: New Computer Revolutionizes Writing. JAMA 204 149-154, April 1, 1968.

22. Glaser, E. L. and Corbato, F. J.: Introduction to Time-Sharing Datamation, pages 24-27, November, 1964.

23. Schwartz, Jules I.: The SDC Time-Sharing System - Part I. Datamation, pages 28-31, November 1964.

24. Schwartz, Jules I.: The SDC Time-Sharing System - Part II. Datamation, pages 51-55, December, 1964.

25. Shaw, J. C.: The Joss System. Datamation, pages 32-36, November 1964.

26. A Panel Discussion on Time-Sharing, Datamation, pages 38-44, November 1964.

27. Harris, D. P.: The PDP-6 Time-Sharing Hardware. Datamation, pages 51-53, November 1964.

28. Lewin, Morton H.: An Introduction to Computer Graphic Terminals. Proc. of the IEEE, 55: 1544-1552, Sept. 1967.

29. Gurley, B. M. and Woodward, C. E.: Light-pen Links Computer Operator. Electronics, pp. 85-87, November 10, 1959.

30. Boole, George: An Investigation of the Laws of Thought. Dover Publications, Inc. 1854.

31. Whitehead, Alfred North and Russell, Bertrand: Principia Mathematica, Volume I. Cambridge University Press. 1950.

32. Ledley, Robert S.: Digital Computational Methods in Symbolic Logic with Examples in Biochemistry. Proc. Natl. Acad. Sci. U. S., 41: 7-42, 1955.

33. Ledley, Robert S. and Lusted, Lee B.: Reasoning Foundations of Medical Diagnosis. Science, 139: 9-21, 3 July 1959.

34. Shannon, Claude E.: A Symbolic Analysis of Relay and Switching Circuits. AIEE Transactions, 57: 713-722, 1938.

35. McCulloch, Warren S. and Pitts, Walter: A Logical Calculus of the Ideas Immanent in Nervous Activity. Bulletin of Mathematical Biophysics, 5: 115-133, 1943.

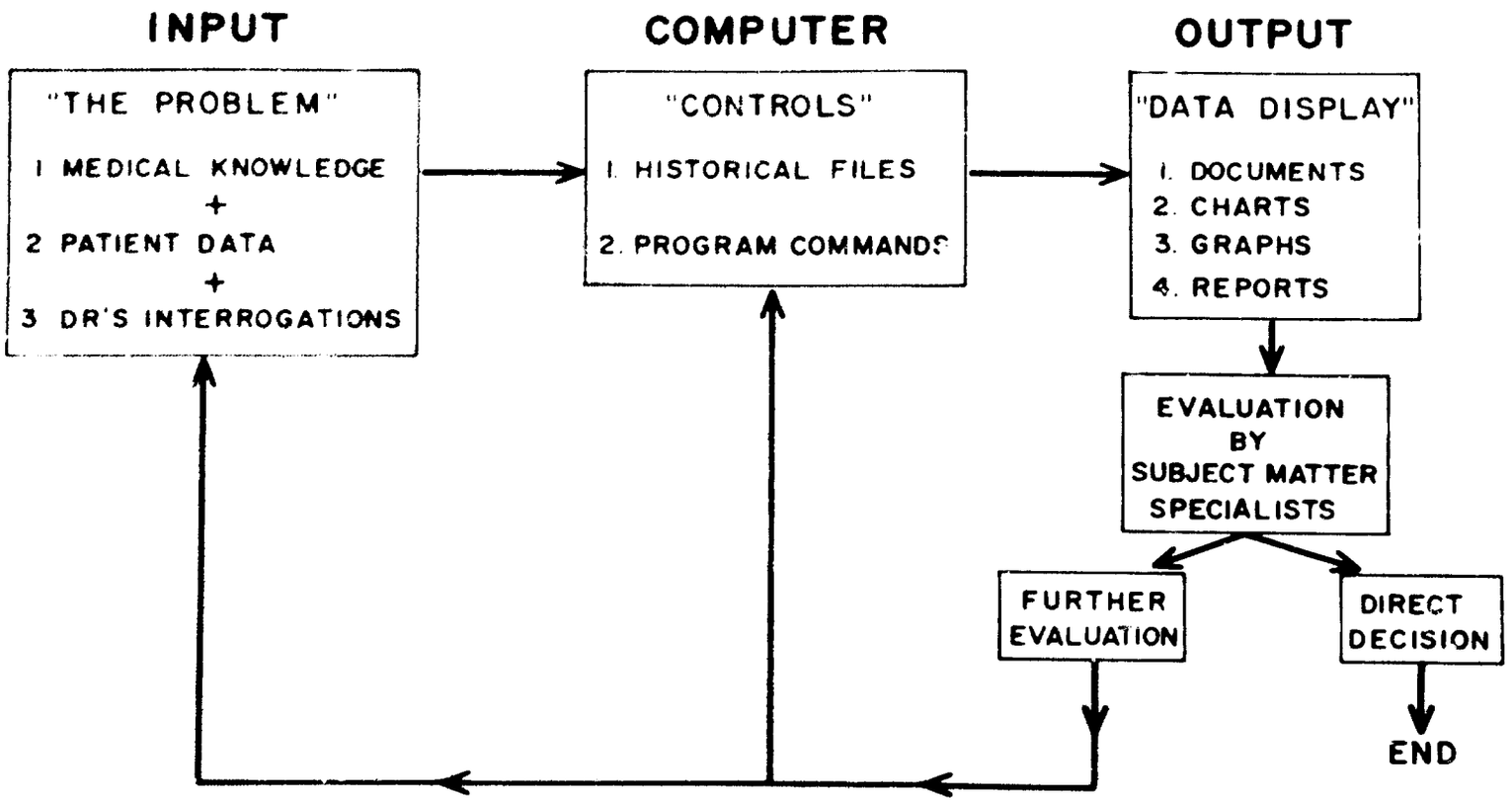
36. McCulloch, Warren S.: The Brain as a Computing Machine. Electrical Engineering, 42: 492-497, June 1949.

37. Bellman, Richard: Dynamic Programming. Science, 153: 34-37, 1 July 1966.

38. Anedecor, George W.: Statistical Methods. The Iowa State College Press, Fourth Edition, 1950.

39. Smith, James G. and Duncan, Acheson J.: Elementary Statistics and Applications. McGraw-Hill Book Company, Inc., New York and London, 1944.

FIGURE I - SCHEMATIC MODEL OF INFORMATION SYSTEM



BIOMEDICAL COMPUTER SYSTEM

FIGURE II

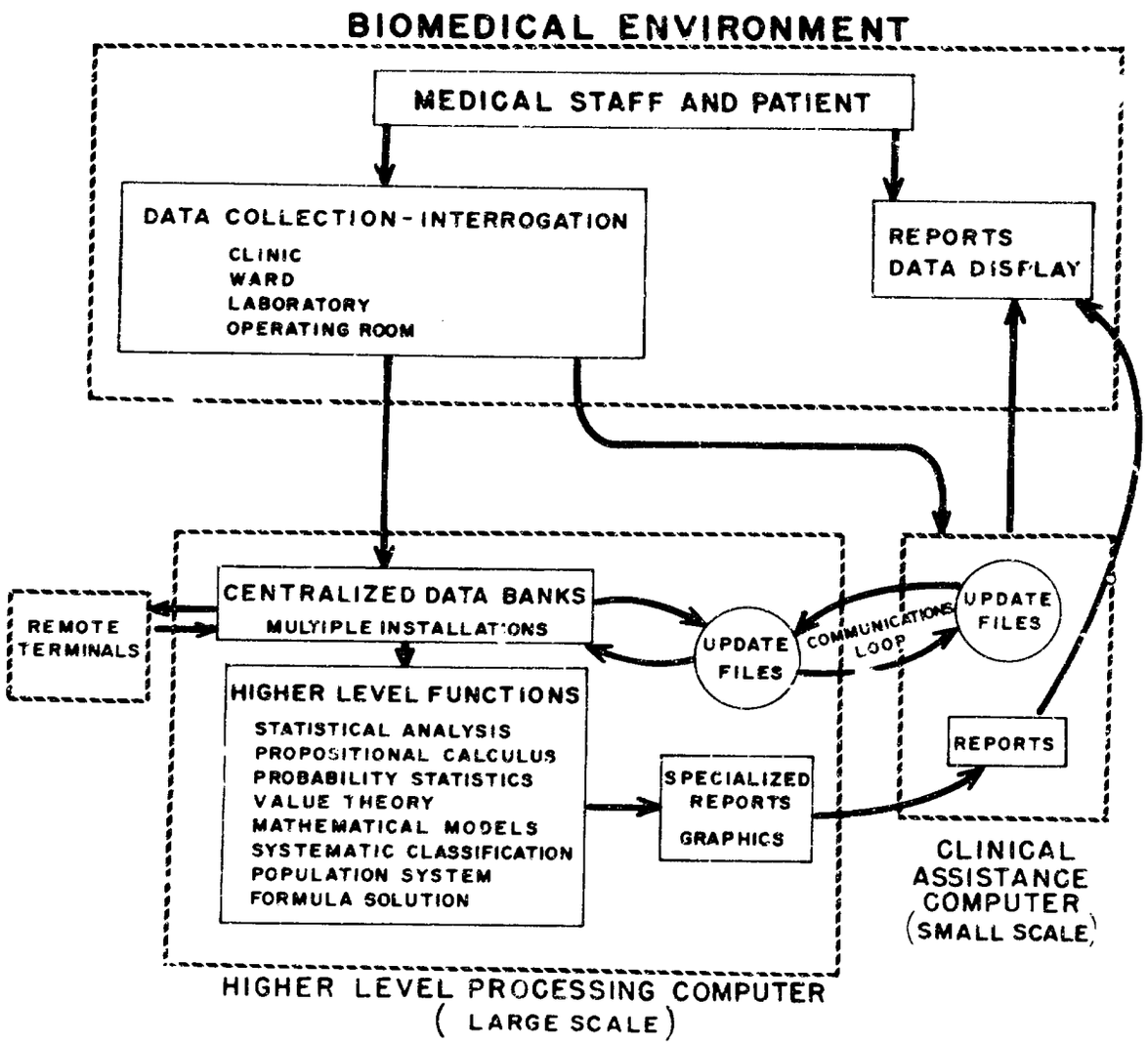


FIGURE III REMOTE CONTROL DATA TRANSMISSION-RECEIVING FOR DIGITAL COMPUTER SYSTEM

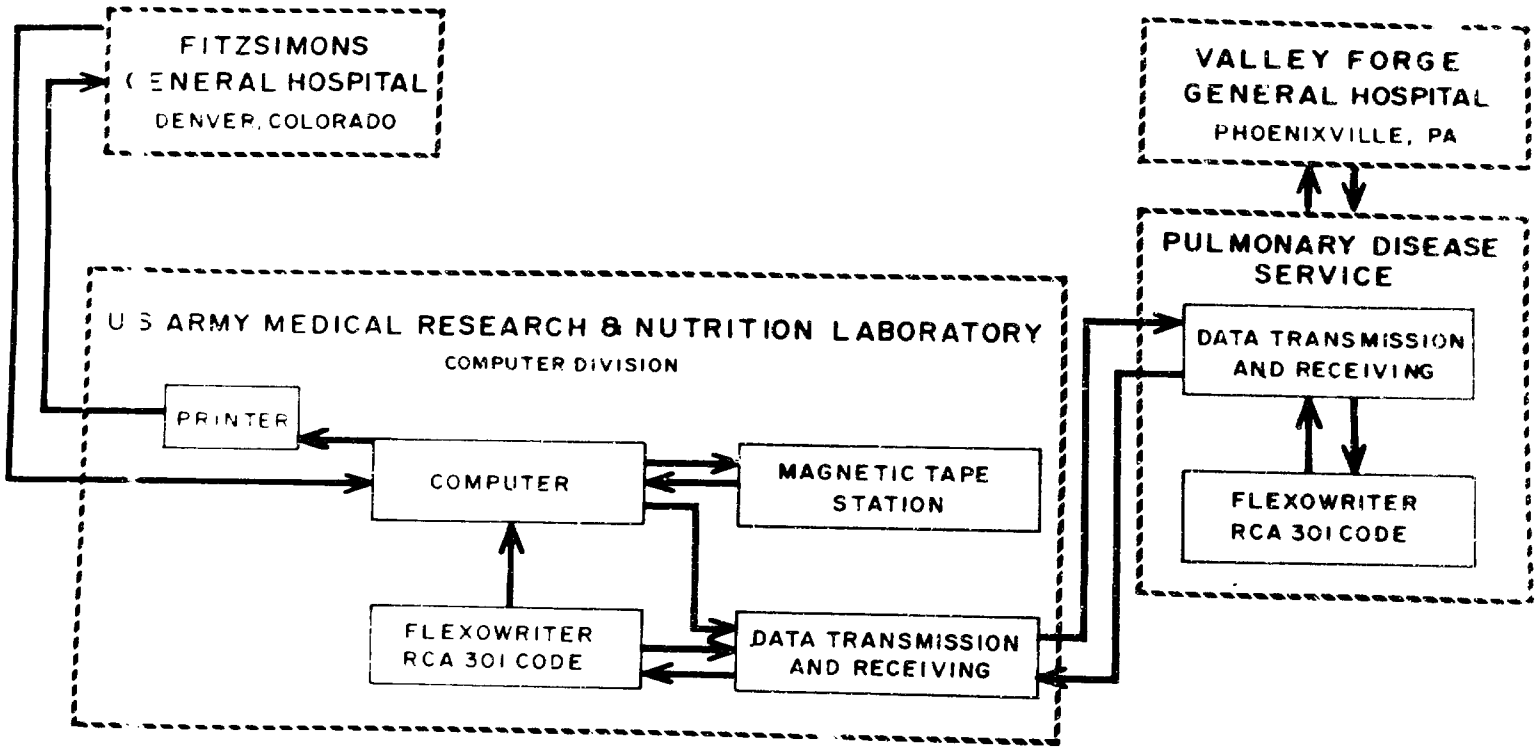


FIGURE IV SPIROMETRY MASTER FILE

CODE	DATE	SEX	HT		WT	AGE	BSA	FVC	FEV_1		M/V	FRQ	MET	MMEF	MEFR	CLINICAL	CASE
63343021000	64349	M	68	34	146	21	177	3338	2932	88	175	100	041	415	1312	0 0 2 1 1 0J 05	11005
63343031000	63343	M	69	36	159	44	185	2120	1263	60	065	045	134	085	0189	1 0 2 1 3 01 05	11006
63343031000	63343	M	69	36	159	44	185	2141	1276	60	083	040	116	095	0154	0 0 2 2 1 01 05	11007
63343041000	63343	M	67	35	156	28	179	3834	3203	84	088	070	050	406	0590	0 0 2 1 1 01 05	11008
63343041000	65011	M	68	34	144	28	177	3042	2361	78	086	070	038	417	0644	0 0 2 1 1 0H 05	11009
63343051000	63343	F	59	31	142	63	155	1796	1257	70	052	060	066	134	0197	0 0 4 2 1 01 05	11010
63343051000	63343	F	59	31	142	63	155	1931	1392	72	042	045	066	148	0211	1 0 4 1 3 01 05	11011
63344011000	63344	M	71	35	156	26	189	5972	5164	86	174	045	050	619	1145	0 0 1 1 1 01 05	11012
63344021000	63344	M	69	35	139	24	177	5298	4266	81	112	070	056	475	0507	1 0 2 1 1 01 05	11013

FIGURE V THE INFORMATION "ITEM". ESTABLISHES THE STRUCTURAL UNIT OF THE DATA BASES

INFORMATION ITEM

CASE	COUGH	SPUTUM	DYSPNEA	CHEST PAIN	DIFFUSE INSP.-EXP. WHEEZES	DIFFUSE RALES	FVC cc's	MVV L/min.	$\frac{FEV_1}{FVC}$ %	ASO_2 %	MET sec
A	1	1	1	Ø	1	0	3600	40	50	88	1.56
B	0	0	0	0	0	0	4600	150	86	96	0.55
C	1	1	1	1	0	1	> 2400	80	75	75	0.88

FVC = FORCED VITAL CAPACITY

FEV_1 = FORCED EXPIRATORY VOLUME IN ONE SECOND

MVV = MAXIMUM VENTILATED VOLUME (MAXIMUM BREATHING CAPACITY)

ASO_2 = ARTERIAL OXYGEN SATURATION

FIGURE VI EXPRESSING QUANTITATIVE DATA AS NON-OVERLAPPING INTERVALS

	INTERVALS																			
	FVC					MVV					FEV_1/FVC					ASO_2				
CASE	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
A	ø	1	ø	ø	ø	ø	ø	ø	ø	1	ø	ø	ø	1	ø	ø	1	ø	ø	ø
B	1	ø	ø	ø	ø	1	ø	ø	ø	ø	1	ø	ø	ø	ø	1	ø	ø	ø	ø
C	ø	ø	ø	1	ø	ø	ø	1	ø	ø	1	ø	ø	ø	ø	ø	ø	1	ø	ø

WHERE FVC

1 = > 5500 - 3900 cc
2 = 3899 - 3100 cc
3 = 3099 - 2600 cc
4 = 2599 - 2100 cc
5 = < 2099 cc

MEASUREMENTS FOR THE MVV, FEV_1/FVC AND ASO_2 WOULD BE SIMILARILY PARTITIONED INTO NON-OVER LAPPING INTERVALS

FIGURE VII THE ROLE OF THE INFORMATION ITEM IN UTILIZING MODERN SYMBOLIC LOGIC

ITEM	SIMPLE STATEMENT	TRUTH VALUE			SYMBOL			NUMERIC VALUE		
		A	B	C	A	B	C	A	B	C
COUGH	COUGH IS PRESENT	T	F	T	S_1	$\bar{S}_1$	S_1	1	0	1
SPUTUM	SPUTUM IS PRESENT	T	F	T	S_2	$\bar{S}_2$	S_2	1	0	1
DYSPNEA	DYSPNEA IS PRESENT	T	F	T	S_3	$\bar{S}_3$	S_3	1	0	1
CHEST PAIN	CHEST PAIN IS PRESENT	F	F	T	$\bar{S}_4$	$\bar{S}_4$	S_4	0	0	1
WHEEZES	WHEEZING IS PRESENT	T	F	F	S_5	$\bar{S}_5$	$\bar{S}_5$	1	0	0
RALES	RALES ARE PRESENT	F	F	T	$\bar{S}_6$	$\bar{S}_6$	S_6	0	0	1
FVC	FVC IS NORMAL	F	T	F	$\bar{S}_7$	S_7	$\bar{S}_7$	0	1	0
MVV	MBC IS NORMAL	F	T	F	$\bar{S}_8$	S_8	$\bar{S}_8$	0	1	0
FEV/FVC	FEV/FVC IS NORMAL	F	T	F	$\bar{S}_9$	S_9	$\bar{S}_9$	0	1	0
ASO_2	ASO_2 IS NORMAL	F	T	F	$\bar{S}_{10}$	S_{10}	$\bar{S}_{10}$	0	1	0
MET	MET IS NORMAL	F	T	F	$\bar{S}_{11}$	S_{11}	$\bar{S}_{11}$	0	1	0

THE SYMBOLIC NOTATION FOR THE NEGATIVE CONDITION UTILIZES THE BAR $(-)$ PLACED IN A HORIZONTAL POS TION OVER THE LETTER $(\bar{S}_1)$ AND IS READ AS IS NOT

FIG. VIII MEDICAL PROBLEM SOLVING INTERACTS WITH THE COMPUTER PROGRAM LIBRARY

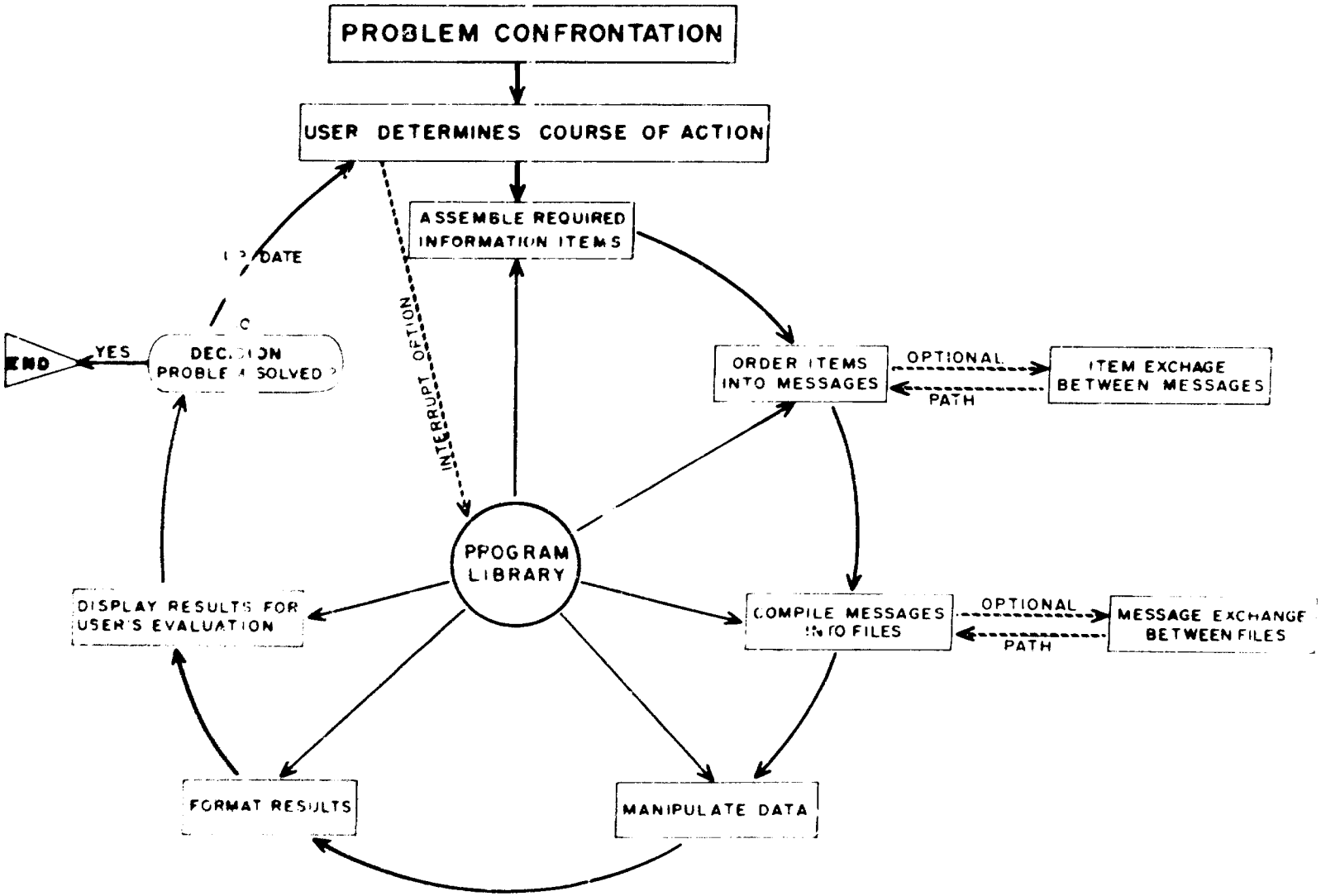
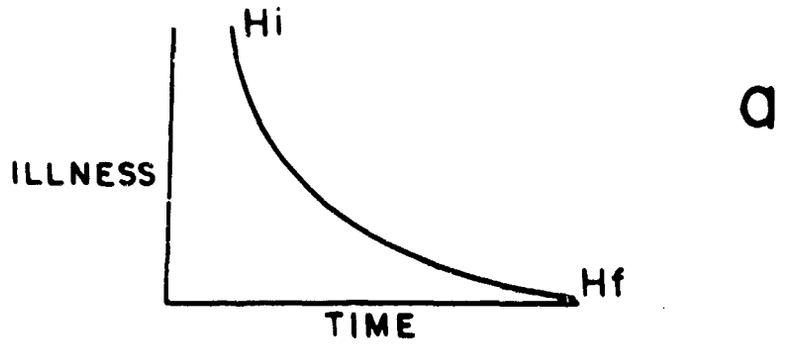
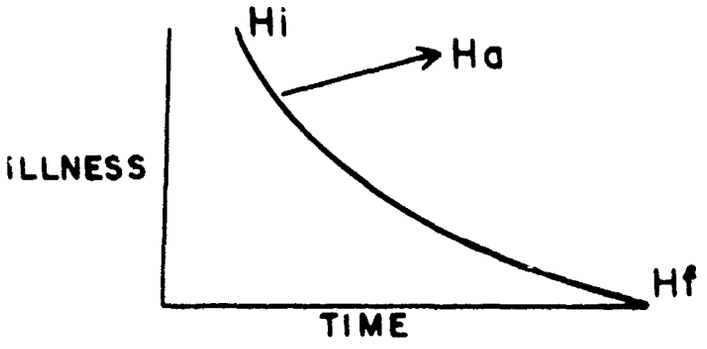


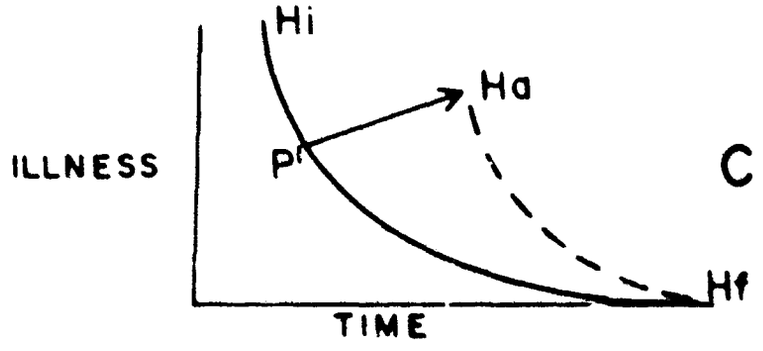
FIGURE IX VARIATIONS OF HEALTH STATE FROM A PROJECTED TRAJECTORY OF A DISEASE PROCESS IN THE NATURAL COURSE OF ILLNESS



a



b



c

FIG. X SPIROMETRY RAW DATA INPUT

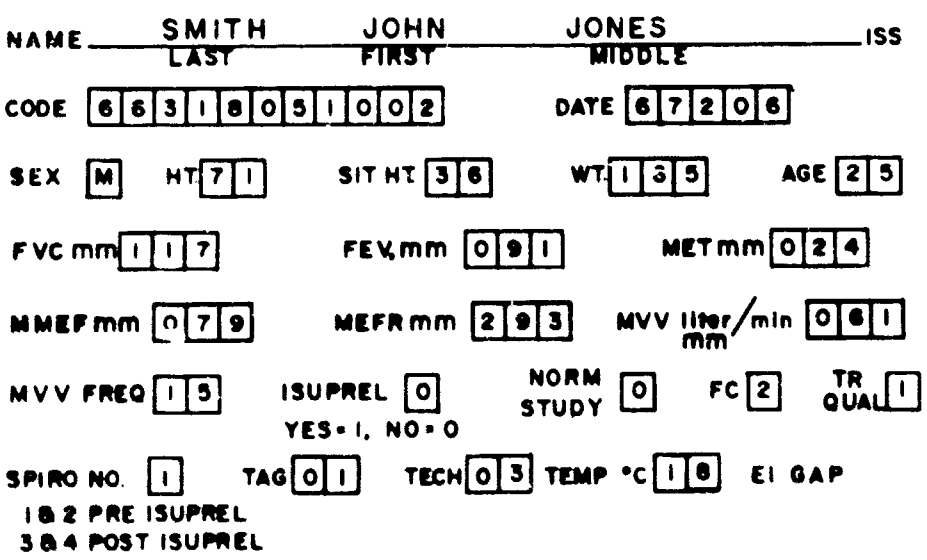


Fig. XI SPIROMETRY CLINICAL REPORT FORM

PULMONARY FUNCTION LABORATORY FITZSIMONS GENERAL HOSPITAL

SPIROMETRY 21 JUN 68

SMITH JOHN JONES 61147011002

AGE: 41 MALE HEIGHT: 71 INCHES WEIGHT: 160 LBS ISUPREL NO

	OBSERVED	PREDICTED	% OF PRFDICTED
FVC	5.223 LITERS	4.927 LITERS	106
FEV_1	4.152 LITERS	4.022 LITERS	103
FEV_1 & OBSV	79%	82%	097
MET	0.63 SECONDS	0.56 SECONDS	111
MMEF	4.06 L/SEC	4.58 L/SEC	089
MEFR	06.96 L/SEC	07.84 L/SEC	089
MVV	173 L/MIN	172 L/MIN	101
MVV FREQ	095/MIN		

Fig XII STATISTICAL ANALYSIS OF SPIROMETRY EXAMINATION

FITZSIMONS GENERAL HOSPITAL PULMONARY FUNCTION CLINIC

DATE OF THIS REPORT: 24 JUN 68 STATISTICAL REPORT OF PULMONARY FUNCTION

NAME: SMITH JOHN JONES CODE: 61147011002

DATE OF SPIROMETRY EXAMINATION: 21 JUN 68

ISUPREL ADMINISTERED: NO TRACING QUALITY: GOOD

STATISTICAL ANALYSIS OF THESE STUDIES DEMONSTRATES A NORMAL AIR VOLUME FUNCTION REFLECTED IN THE VITAL CAPACITY (FVC) PERFORMANCE AND A NORMAL AIR FLOW FUNCTION REFLECTED IN THE NUMERICAL VALUE OF THE RATIO FEV-1: FVC.

THE STATISTICAL CRITERIA FOR THIS INDIVIDUAL ARE BASED UPON COMPUTED PREDICTED VALUES DERIVED FROM A NORMAL POPULATION.

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
U. S. Army Medical Research & Nutrition Laboratory Fitzsimons General Hospital Denver, Colorado 80240	Unclassified 2b. GROUP

3. REPORT TITLE

A COMPUTER BASED BIOMEDICAL INFORMATION SYSTEM
I. LOGIC FOUNDATIONS AND TECHNIQUES

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Laboratory Report Number 320

5. AUTHOR(S) (First name, middle initial, last name)

James C. Syner

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
October 1968	50	39

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO. 3A025601A822 Mil. Int. Med. 067 Computer Class.of Pulmonary Disability c. 064 Biomed.Information Systems Design 068 Computer Instrument Linkage 081 Digital Comp.based Biomed.Infor.System d. to support special forces troop in the field	USAMRNL Report No. 320 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT Distribution of this document is unlimited.

[redacted] 60240

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
USAMRNL Laboratory Report	U.S. Army Medical Research and Development Command, Washington, D. C.

13. ABSTRACT

A digital computer based biomedical information system has been designed to service the needs of physicians engaged in patient care and clinical research, and scientists engaged in laboratory research. The system embraces all functions of information processing which include information collection, storage, retrieval, analyses and display. The principal goal of the project is to place these functions under the maximum degree of automation possible with existing hardware-software capabilities. At the time of this report the status of the system is best characterized as "semi-automatic." From the time of inception, and throughout implementation, the project has been carried out under the concepts, principles and techniques of the systems analyses. Experience has demonstrated that a total biomedical information system to service a complex medical facility such as the Fitzsimons General Hospital establishment will necessarily evolve through several generations of hardware-software alterations. There will probably never be a total solution to the problem but rather sequential events viewed as minor to major achievements representing progress towards an ever improving system. Progress is absolutely dependent upon support and participation from the professional staff, including both physicians and scientific investigators.

DD FORM 1 NOV 68 1473 REPLACES DD FORM 1473, 1 JAN 64, WHICH IS OBSOLETE FOR ARMY USE.

UNCLASSIFIED
Security Classification

UNCLASSIFIED
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Digital Computer						
Analog Computer						
Information System						
Computer Programs						
Systems Analysis						
Systems Design						
Hospital						
Laboratory						

UNCLASSIFIED
Security Classification